



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C07H 21/02, 2/04, C12N 5/00, 5/10, 15/00, 15/09, 15/11, 15/31	A1	(11) International Publication Number: WO 98/58943 (43) International Publication Date: 30 December 1998 (30.12.98)
(21) International Application Number: PCT/US98/12764 (22) International Filing Date: 18 June 1998 (18.06.98) (30) Priority Data: 60/050,359 20 June 1997 (20.06.97) US 60/053,377 22 July 1997 (22.07.97) US 60/053,344 22 July 1997 (22.07.97) US 60/057,483 3 September 1997 (03.09.97) US (71) Applicants (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US). MEDIMMUNE, INC. [US/US]; 35 West Watkins Mill Road, Gaithersburg, MD 20878 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): FRASER, Claire [US/US]; 11915 Glen Mill Road, Potomac, MD 20854 (US). WHITE, Owen, R. [US/US]; 886 Quince Orchard Boulevard #202, Gaithersburg, MD 20878 (US). CLAYTON, Rebecca [US/US]; 6706 B. Polor Avenue, Takoma Park, MD 20912 (US). DOUGHERTY, Brian, A. [US/US]; 10 Rosemary Lane, Killingworth, CT 06419 (US). LATHIGRA, Raju [IN/US]; 19051 Steeple Place, Germantown, MD 20874		(US). SMITH, Hamilton, O. [US/US]; 8222 Carrbridge Circle, Towson, MD 21204 (US). (74) Agents: HOOVER, Kenley, K. et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: <i>BORRELIA BURGDORFERI</i> POLYNUCLEOTIDES AND SEQUENCES		
(57) Abstract <p>The present invention provides polynucleotide sequences of the genome of <i>Borrelia burgdorferi</i>, polypeptide sequences encoded by the polynucleotide sequences, corresponding polynucleotides and polypeptides, vectors and hosts comprising the polynucleotides, and assays and other uses thereof. The present invention further provides polynucleotide and polypeptide sequence information stored on computer readable media, and computer-based systems and methods which facilitate its use.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

***Borrelia burgdorferi* Polynucleotides and Sequences**

5

Field of the Invention

The present invention relates to the field of molecular biology. In particular, it relates to, among other things, nucleotide sequences of *Borrelia burgdorferi*, contigs, ORFs, fragments, probes, primers and related polynucleotides thereof, peptides and polypeptides encoded by the sequences, and uses of the polynucleotides and sequences thereof, such as in fermentation, polypeptide production, assays and pharmaceutical development, among others.

Statement as to Rights to Inventions Made Under Federally-Sponsored Research and Development

Part of the work performed during development of this invention utilized U.S. Government funds. The U.S. Government may have certain rights in the invention - DE-FC02-95ER61962; DE-FC02-95ER61963; and NAGW 2554.

20

Background of the Invention

Spirochetes are a family of motile, unicellular, spiral-shaped bacteria which share a number of structural characteristics. Three genera of the spirochetes are pathogenic in humans: (a) *Treponema*, which includes the pathogens that cause syphilis (*T. pallidum*), yaws (*T. pertenue*), and pinta (*T. carateum*); (b) *Borrelia*, which includes the pathogens that cause epidemic and endemic relapsing fever and Lyme disease; and (c) *Leptospira*, which includes a wide variety of small spirochetes that cause mild to serious systemic human illness (Koff, A. B. and Rosen, T. *J. Am. Acad. Dermatol.* **29**:519-535 (1993)).

Lyme borreliosis, more commonly known as Lyme disease, is presently the most common human disease in the United States transmitted by an arthropod vector. Centers for Disease Control, Morbid. Mortal. Weekly Rep. 44:590-591 (1995). Further, infection of household pets, such as dogs, is a considerable problem. The causative agent of this affliction is the spirochete *Borrelia burgdorferi*, which is generally transmitted to mammalian hosts by feeding ticks. Barbour, A. and Fish, D. *Science* 260:1610-1616 (1993). Once the bacteria pass through the skin they disseminate and produce a variety of clinical manifestations. Diagnosis of this disease is often made serologically by the identification of antiborrelial antibodies. Hilton, E. et al., *J. Clin. Microbiol.* 35:774-776 (1997).

While initial symptoms often include a rash at the infection point, Lyme disease is a multisystemic disorder that may include arthritic, carditic, and neurological manifestations. While antibiotics are currently used to treat active cases of Lyme disease, *B. burgdorferi* appears to be able to persist even after prolonged antibiotic treatment. Further, *B. burgdorferi* can persist for years in a mammalian host even in the presence of an active immune response. Straubinger, R. et al., J. Clin. Microbiol. 35:111-116 (1997); Steere, A., N. Engl. J. Med. 321:586-596 (1989).

Animal models have proven useful for studying the progression of Lyme disease, methods for preventing this disease, and immunological responses to antigenic challenges with *B. burgdorferi* proteins. Garcia-Monoco, J. et al., J. Infect. Dis. 175:1243-1245 (1997). Using a canine model, Starubinger, R. et al., Infect. Immun. 65:1273-1285 (1977), demonstrated that *B. burgdorferi* migrates into joints and induces up-regulation of interleukin-8 in synovial membranes. Similarly, *B. burgdorferi* induction of interleukin-8 production has been demonstrated in cultured human endothelial cells. Burns, M. et al., Infect. Immun. 65:1217-1222 (1997).

Antigenic heterogeneity has been postulated as a mechanism used by *B. burgdorferi* for evasion of host immune responses. Schwan, T. et al., Can. J. Microbiol. 37:450-454 (1991). In support of this mechanism, antigenic variation has been described with other pathogenic bacteria. Hagbloom, P. et al., Nature 315:156-158 (1985). Further, cassette type genetic recombination of genes encoding *B. burgdorferi* surface proteins has been shown to decrease the antigenicity of these organisms to antibodies generated against strains which have not undergone the same recombination. Zhang, J. et al., Cell 89:275-285 (1997).

A number of different types of Lyme disease vaccines have been tested and shown to induce immunological responses. Whole-cell *B. burgdorferi* vaccines have been shown to induce both immunological responses and protective immunity in several animal models. Reviewed in Wormser, G., Clin. Infect. Dis. 21:1267-1274 (1995). For example, dogs inoculated with a chemically inactivated whole-cell vaccine primarily develop antibodies to outer surface membrane proteins of the administered organism. Further, passive immunity has been also demonstrated in animals using *B. burgdorferi* specific antisera. Similarly, passive immunity is conferred human by the administration of sera obtained from Lyme disease patients.

While whole-cell Lyme disease vaccines confer protective immunity in animal models, use of such vaccines presents the risk that responsive antibodies will be generated which cross react with human antigens. Reviewed in Wormser, G., supra. This problem is at least partly the result of the production of *B. burgdorferi* specific antibodies which cross-react with hepatocytes and both muscle and nerve cells. *B. burgdorferi* heat shock proteins and the 41-kd flagellin subunit are believed to contain the antigens against which these cross-reactive antibodies are generated.

It is clear that the etiology of diseases mediated or exacerbated by *B. burgdorferi* genes, and that characterizing the genes and their patterns of expression would add dramatically to our

understanding of the organism and its host interactions. Knowledge of *B. burgdorferi* genes and genomic organization would dramatically improve understanding of disease etiology and lead to improved and new ways of preventing, ameliorating, arresting and reversing diseases. Moreover, characterized genes and genomic fragments of *B. burgdorferi* would provide reagents for, among other things, detecting, characterizing and controlling *B. burgdorferi* infections. There is a need therefore to characterize the genome of *B. burgdorferi* and for polynucleotides and sequences of this organism.

SUMMARY OF THE INVENTION

The present invention is based on the sequencing of fragments of the *Borrelia burgdorferi* genome. The primary nucleotide sequences which were generated are provided in SEQ ID NOS:1-155.

The present invention provides the complete nucleotide sequence of the *Borrelia burgdorferi* chromosome and 154 contigs representing the majority of the sequence of the *B. burgdorferi* extrachromosomal elements, all of which are listed in tables below and set out in the Sequence Listing submitted herewith, and representative fragments thereof, in a form which can be readily used, analyzed, and interpreted by a skilled artisan. In one embodiment, the present invention is provided as contiguous strings of primary sequence information corresponding to the nucleotide sequences depicted in SEQ ID NOS: 1-155.

The present invention further provides nucleotide sequences which are at least 95%, 96%, 97%, 98%, and 99%, identical to the nucleotide sequences of SEQ ID NOS:1-155, ORF IDs and corresponding ORFs.

The nucleotide sequences of SEQ ID NOS:1-155, ORF ID or ORF within, a representative fragment thereof, or a nucleotide sequence which is at least 95% identical to said nucleotide sequence may be provided in a variety of mediums to facilitate its use. In one application of this embodiment, the sequences of the present invention are recorded on computer readable media. Such media includes, but is not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media.

The present invention further provides systems, particularly computer-based systems which contain the sequence information herein described stored in a data storage means. Such systems are designed to identify commercially important fragments of the *Borrelia burgdorferi* genome.

Another embodiment of the present invention is directed to fragments of the *Borrelia burgdorferi* genome having particular structural or functional attributes. Such fragments of the *Borrelia burgdorferi* genome of the present invention include, but are not limited to, fragments which encode peptides, hereinafter referred to as open reading frames or ORFs, fragments which modulate the expression of an operably linked ORF, hereinafter referred to as expression

modulating fragments or EMFs, and fragments which can be used to diagnose the presence of *Borrelia burgdorferi* in a sample, hereinafter referred to as diagnostic fragments or DFs.

Each of the ORF IDs and ORFs in fragments of the *Borrelia burgdorferi* genome disclosed in Tables 1-6, and the EMFs found 5' prime of the initiation codon, can be used in numerous ways as polynucleotide reagents. For instance, the sequences can be used as diagnostic probes or amplification primers for detecting or determining the presence of a specific microbe in a sample, to selectively control gene expression in a host and in the production of polypeptides, such as polypeptides encoded by ORFs of the present invention, particular those polypeptides that have a pharmacological activity.

The present invention further includes recombinant constructs comprising one or more fragments of the *Borrelia burgdorferi* genome of the present invention. The recombinant constructs of the present invention comprise vectors, such as a plasmid or viral vector, into which a fragment of the *Borrelia burgdorferi* has been inserted.

The present invention further provides host cells containing any of the isolated fragments of the *Borrelia burgdorferi* genome of the present invention. The host cells can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic cell, such as a yeast cell, or a procaryotic cell such as a bacterial cell.

The present invention is further directed to isolated polypeptides and proteins encoded by ORFs of the present invention. A variety of methods, well known to those of skill in the art, routinely may be utilized to obtain any of the polypeptides and proteins of the present invention. For instance, polypeptides and proteins of the present invention having relatively short, simple amino acid sequences readily can be synthesized using commercially available automated peptide synthesizers. Polypeptides and proteins of the present invention also may be purified from bacterial cells which naturally produce the protein. Yet another alternative is to purify polypeptide and proteins of the present invention from cells which have been altered to express them.

The invention further provides methods of obtaining homologs of the fragments of the *Borrelia burgdorferi* genome of the present invention and homologs of the proteins encoded by the ORFs of the present invention. Specifically, by using the nucleotide and amino acid sequences disclosed herein as a probe or as primers, and techniques such as PCR cloning and colony/plaque hybridization, one skilled in the art can obtain homologs.

The invention further provides antibodies which selectively bind polypeptides and proteins of the present invention. Such antibodies include both monoclonal and polyclonal antibodies.

The invention further provides hybridomas which produce the above-described antibodies. A hybridoma is an immortalized cell line which is capable of secreting a specific monoclonal antibody.

The present invention further provides methods of identifying test samples derived from cells which express one of the ORFs of the present invention, or a homolog thereof. Such

methods comprise incubating a test sample with one or more of the antibodies of the present invention, or one or more of the DFs of the present invention, under conditions which allow a skilled artisan to determine if the sample contains the ORF or product produced therefrom.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the above-described assays.

Specifically, the invention provides a compartmentalized kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the antibodies, or one of the DFs of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of bound antibodies or hybridized DFs.

Using the isolated proteins of the present invention, the present invention further provides methods of obtaining and identifying agents capable of binding to a polypeptide or protein encoded by one of the ORFs of the present invention. Specifically, such agents include, as further described below, antibodies, peptides, carbohydrates, pharmaceutical agents and the like. Such methods comprise steps of: (a) contacting an agent with an isolated protein encoded by one of the ORFs of the present invention; and (b) determining whether the agent binds to said protein.

The present genomic sequences of *Borrelia burgdorferi* will be of great value to all laboratories working with this organism and for a variety of commercial purposes. Many fragments of the *Borrelia burgdorferi* genome will be immediately identified by similarity searches against GenBank or protein databases and will be of immediate value to *Borrelia burgdorferi* researchers and for immediate commercial value for the production of proteins or to control gene expression.

The methodology and technology for elucidating extensive genomic sequences of bacterial and other genomes has and will greatly enhance the ability to analyze and understand chromosomal organization. In particular, sequenced contigs and genomes will provide the models for developing tools for the analysis of chromosome structure and function, including the ability to identify genes within large segments of genomic DNA, the structure, position, and spacing of regulatory elements, the identification of genes with potential industrial applications, and the ability to do comparative genomic and molecular phylogeny.

DESCRIPTION OF THE FIGURES

FIGURE 1 is a block diagram of a computer system (102) that can be used to implement computer-based systems of present invention.

FIGURE 2 is a schematic diagram depicting the data flow and computer programs used to collect, assemble, edit and annotate the contigs of the *Borrelia burgdorferi* genome of the present invention. Both Macintosh and Unix platforms are used to handle the AB 373 and 377 sequence data files, largely as described in Kerlavage *et al.*, *Proceedings of the Twenty-Sixth*

Annual Hawaii International Conference on System Sciences, 585, IEEE Computer Society Press, Washington D.C. (1993). Factura (AB) is a Macintosh program designed for automatic vector sequence removal and end-trimming of sequence files. The program Loadis runs on a Macintosh platform and parses the feature data extracted from the sequence files by Factura to the
5 Unix based *Borrelia burgdorferi* relational database. Assembly of contigs (and whole genome sequences) is accomplished by retrieving a specific set of sequence files and their associated features using Extrseq, a Unix utility for retrieving sequences from an SQL database. The resulting sequence file is processed to trim portions of the sequences with a high rate ambiguous nucleotides. The sequence files were assembled using TIGR Assembler, an assembly engine
10 designed at The Institute for Genomic Research (TIGR) for rapid and accurate assembly of thousands of sequence fragments. The collection of contigs generated by the assembly step is loaded into the database with the lassie program. Identification of open reading frames (ORFs) is accomplished by processing contigs with zorf. The ORFs are searched against *B. burgdorferi* sequences from GenBank and against all protein sequences using the BLASTN and BLASTP
15 programs, described in Altschul *et al.*, *J. Mol. Biol.* 215: 403-410 (1990). Results of the ORF determination and similarity searching steps were loaded into the database. As described below, some results of the determination and the searches are set out in Tables 1-6.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

20 The present invention is based on the sequencing of fragments of the *Borrelia burgdorferi* genome and analysis of the sequences. The primary nucleotide sequences generated by sequencing the fragments are provided in SEQ ID NOS: 1-155. (As used herein, the "primary sequence" refers to the nucleotide sequence represented by the IUPAC nomenclature system.)
SEQ ID NOS:1-155

25 In addition, the present invention provides the nucleotide sequences of SEQ ID NOS: 1-155, or representative fragments thereof, in a form which can be readily used, analyzed, and interpreted by a skilled artisan.

As used herein, a "representative fragment of the nucleotide sequence depicted in SEQ ID NOS:1-155" refers to any portion of the SEQ ID NOS: 1-155 which is not presently
30 represented within a publicly available database. Preferred representative fragments of the present invention are *Borrelia burgdorferi* open reading frames (ORFs) represented by ORF IDs, expression modulating fragments (EMFs) and diagnostic fragments (DFs) which can be used to diagnose the presence of *Borrelia burgdorferi* in sample. A non-limiting identification of preferred representative portions are provided in Tables 1-6 as ORF IDs. As discussed in detail
35 below, the information provided in SEQ ID NOS:1-155 and in Tables 1-6 together with routine cloning, synthesis, sequencing and assay methods will enable those skilled in the art to clone and sequence all "representative fragments" of interest, including ORFs encoding a large variety of *Borrelia burgdorferi* proteins.

The present invention is further directed to nucleic acid molecules encoding portions or fragments of the nucleotide sequences described herein. Fragments include portions of the nucleotide sequences of Table 1-6 (ORF IDs) and SEQ ID NOS:1-155, at least 10 contiguous nucleotides in length selected from any two integers, one of which representing a 5' nucleotide position and a second of which representing a 3' nucleotide position, where the first nucleotide for each nucleotide sequence in SEQ ID NOS:1-155 is position 1 (therefore, the sequence positions for each ORF ID is determined by the numbering of the SEQ ID comprising the ORF ID). That is, every combination of a 5' and 3' nucleotide position that a fragment at least 10 contiguous nucleotides in length could occupy is included in the invention. At least means a fragment may be 10 contiguous nucleotide bases in length or any integer between 10 and the length of an entire nucleotide sequence of SEQ ID NOS:1-155 minus 1. Therefore, included in the invention are contiguous fragments specified by any 5' and 3' nucleotide base positions of a nucleotide sequences of SEQ ID NOS:1-155 wherein the contiguous fragment is any integer between 10 and the length of an entire nucleotide sequence minus 1.

Further, the invention includes polynucleotides comprising fragments specified by size, in nucleotides, rather than by nucleotide positions. The invention includes any fragment size, in contiguous nucleotides, selected from integers between 10 and the length of an entire ORF ID or SEQ ID NO., minus 1. Preferred sizes of contiguous nucleotide fragments include 20 nucleotides, 30 nucleotides, 40 nucleotides, 50 nucleotides. Other preferred sizes of contiguous nucleotide fragments, which may be useful as diagnostic probes and primers, include fragments 50-300 nucleotides in length which include, as discussed above, fragment sizes representing each integer between 50-300. Larger fragments are also useful according to the present invention corresponding to most, if not all, of the nucleotide sequences shown in Tables 1-6 (ORF IDs) and SEQ ID NOS:1-155. The preferred sizes are, of course, meant to exemplify not limit the present invention as all size fragments, representing any integer between 10 and the length of an entire nucleotide sequence minus 1, of each ORF ID and SEQ ID NO., are included in the invention.

The present invention also provides for the exclusion of any fragment, specified by 5' and 3' base positions or by size in nucleotide bases as described above for any ORF ID or SEQ ID NOS:1-155. Any number of fragments of nucleotide sequences in ORF IDs or SEQ ID NOS:1-155, specified by 5' and 3' base positions or by size in nucleotides, as described above, may be excluded from the present invention.

While the presently disclosed sequences of SEQ ID NOS: 1-155 are highly accurate, sequencing techniques are not perfect and, in relatively rare instances, further investigation of a fragment or sequence of the invention may reveal a nucleotide sequence error present in a nucleotide sequence disclosed in SEQ ID NOS: 1-155. However, once the present invention is made available (*i.e.*, once the information in SEQ ID NOS: 1-155 and Tables 1-6 has been made available), resolving a rare sequencing error in SEQ ID NOS: 1-155 will be well within the skill

of the art. The present disclosure makes available sufficient sequence information to allow any of the described contigs or portions thereof to be obtained readily by straightforward application of routine techniques. Further sequencing of such polynucleotide may proceed in like manner using manual and automated sequencing methods which are employed ubiquitous in the art. Nucleotide sequence editing software is publicly available. For example, Applied Biosystem's (AB) AutoAssembler can be used as an aid during visual inspection of nucleotide sequences. By employing such routine techniques potential errors readily may be identified and the correct sequence then may be ascertained by targeting further sequencing effort, also of a routine nature, to the region containing the potential error.

Even if all of the very rare sequencing errors in SEQ ID NOS: 1-155 were corrected, the resulting nucleotide sequences would still be at least 95% identical, nearly all would be at least 99% identical, and the great majority would be at least 99.9% identical to the nucleotide sequences of SEQ ID NOS: 1-155.

As discussed elsewhere herein, polynucleotides of the present invention readily may be obtained by routine application of well known and standard procedures for cloning and sequencing DNA. Detailed methods for obtaining libraries and for sequencing are provided below, for instance. A wide variety of *Borrelia burgdorferi* strains that can be used to prepare *B. burgdorferi* genomic DNA for cloning and for obtaining polynucleotides of the present invention are available to the public from recognized depository institutions, such as the American Type Culture Collection (ATCC). While the present invention is enabled by the sequences and other information herein disclosed, the *B. burgdorferi* strain that provided the DNA of the present Sequence Listing, has been deposited with the ATCC, 10801 University Blvd. Manassas, VA 20110-2209, as Deposit No. 202012, on 8 August 1997. The ATCC Deposit is provided merely as a convenience to those of skill in the art. Reference to the deposit is not a waiver of any rights of the inventors or their assignees in the present subject matter.

The nucleotide sequences of the genomes from different strains of *Borrelia burgdorferi* differ somewhat. However, the nucleotide sequences of the genomes of all *Borrelia burgdorferi* strains will be at least 95% identical, in corresponding part, to the nucleotide sequences provided in SEQ ID NOS: 1-155 and the ORF IDs within. Nearly all will be at least 99% identical and the great majority will be 99.9% identical.

The present application is further directed to nucleic acid molecules at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleic acid sequence shown in SEQ ID NOS: 1-155 and the ORF IDs within. The above nucleic acid sequences are included irrespective of whether they encode a polypeptide having *B. burgdorferi* activity. This is because even where a particular nucleic acid molecule does not encode a polypeptide having *B. burgdorferi* activity, one of skill in the art would still know how to use the nucleic acid molecule, for instance, as a hybridization probe. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having *B. burgdorferi* activity include, *inter alia*, isolating a *B. burgdorferi* gene or allelic variants thereof from a DNA library, and detecting *B. burgdorferi* mRNA expression from

biological or environmental samples, suspected of containing *B. burgdorferi* by Northern Blot, PCR, or similar analysis.

Preferred, are nucleic acid molecules having sequences at least 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in SEQ ID NOS: 1-155, the ORF IDs, and the ORF within each ORF ID, which do, in fact, encode a polypeptide having *B. burgdorferi* protein activity. By "a polypeptide having *B. burgdorferi* activity" is intended polypeptides exhibiting activity similar, but not necessarily identical, to an activity of the *B. burgdorferi* protein of the invention, as measured in a particular biological assay suitable for measuring activity of the specified protein.

Due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the nucleic acid molecules having a sequence at least 90%, 95%, 96%, 97%, 98%, or 99% identical to the nucleic acid sequences shown in SEQ ID NOS: 1-155, the ORF IDs, and the ORF within each ORF ID, will encode a polypeptide having *B. burgdorferi* protein activity. In fact, since degenerate variants of these nucleotide sequences all encode the same polypeptide, this will be clear to the skilled artisan even without performing the above described comparison assay. It will be further recognized in the art that, for such nucleic acid molecules that are not degenerate variants, a reasonable number will also encode a polypeptide having *B. burgdorferi* protein activity. This is because the skilled artisan is fully aware of amino acid substitutions that are either less likely or not likely to significantly effect protein function (e.g., replacing one aliphatic amino acid with a second aliphatic amino acid), as further described below.

The biological activity or function of the polypeptides of the present invention are expected to be similar or identical to polypeptides from other bacteria that share a high degree of structural identity/similarity. Tables 1, 2, 4, and 5 lists accession numbers and descriptions for the closest matching sequences of polypeptides available through Genbank. It is therefore expected that the biological activity or function of the polypeptides of the present invention will be similar or identical to those polypeptides from other bacterial genuses, species, or strains listed in Tables 1, 2, 4, and 5.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the *B. burgdorferi* polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted, inserted, or substituted with another nucleotide. The query sequence may be an entire sequence shown in SEQ ID NOS: 1-155, an ORF ID, or the ORF within each ORF ID, or any fragment specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. *See* Brutlag et al. (1990) *Comp. App. Biosci.* 6:237-245. In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by first converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only nucleotides outside the 5' and 3' nucleotides of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 nucleotide subject sequence is aligned to a 100 nucleotide query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 nucleotides at 5' end. The 10 unpaired nucleotides represent 10% of the sequence (number of nucleotides at the 5' and 3' ends not matched/total number of nucleotides in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 nucleotides were perfectly matched the final percent identity would be 90%. In another example, a 90 nucleotide subject sequence is compared with a 100 nucleotide query sequence. This time the deletions are internal deletions so that there are no nucleotides on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only nucleotides 5' and 3' of the

subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

COMPUTER RELATED EMBODIMENTS

5 The nucleotide sequences provided in SEQ ID NOS: 1-155, including ORF IDs and corresponding ORFs, a representative fragment thereof, or a nucleotide sequence at least 95%, preferably at least 96%, 97%, 98% or 99%, and most preferably at least 99.9% identical to said nucleotide sequences may be "provided" in a variety of mediums to facilitate use thereof. As used herein, provided refers to a manufacture, other than an isolated nucleic acid molecule,
10 which contains a nucleotide sequence of the present invention, a representative fragment thereof, or a nucleotide sequence at least 95%, preferably at least 99% and most preferably at least 99.9% identical to a polynucleotide of the present invention. Such a manufacture provides a large portion of the *Borrelia burgdorferi* genome and parts thereof (e.g., a *Borrelia burgdorferi* open reading frame (ORF)) in a form which allows a skilled artisan to examine the manufacture using
15 means not directly applicable to examining the *Borrelia burgdorferi* genome or a subset thereof as it exists in nature or in purified form.

 In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are
20 not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories, such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a
25 nucleotide sequence of the present invention. Likewise, it will be clear to those of skill how additional computer readable media that may be developed also can be used to create analogous manufactures having recorded thereon a nucleotide sequence of the present invention.

 As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for
30 recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

 A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen
35 to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase,

Oracle, or the like. A skilled artisan can readily adapt any number of data-processor structuring formats (e.g., text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. Thus, by providing in computer readable form the nucleotide sequences of the present invention (e.g. SEQ ID NOS: 1-155), a representative fragment thereof, or a nucleotide sequence at least 95%, preferably at least 96%, 97%, 98%, 99% and most preferably at least 99.9% identical to a sequence of the present invention (e.g. SEQ ID NOS: 1-155) enables the skilled artisan routinely to access the provided sequence information for a wide variety of purposes.

The examples which follow demonstrate how software which implements the BLAST (Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990)) and BLAZE (Brutlag *et al.*, *Comp. Chem.* 17:203-207 (1993)) search algorithms on a Sybase system was used to identify open reading frames (ORFs) within the *Borrelia burgdorferi* genome which contain homology to ORFs or proteins from both *Borrelia burgdorferi* and from other organisms. Among the ORFs discussed herein are protein encoding fragments of the *Borrelia burgdorferi* genome useful in producing commercially important proteins, such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

The present invention further provides systems, particularly computer-based systems, which contain the sequence information described herein. Such systems are designed to identify, among other things, commercially important fragments of the *Borrelia burgdorferi* genome.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based system are suitable for use in the present invention.

As stated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means.

As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the present genomic sequences which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of

commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, MacPattern (EMBL), BLASTN and BLASTX (NCBIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems.

As used herein, a "target sequence" can be any DNA or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzymic active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

A variety of structural formats for the input and output means can be used to input and output the information in the computer-based systems of the present invention. A preferred format for an output means ranks fragments of the *Borrelia burgdorferi* genomic sequences possessing varying degrees of homology to the target sequence or target motif. Such presentation provides a skilled artisan with a ranking of sequences which contain various amounts of the target sequence or target motif and identifies the degree of homology contained in the identified fragment.

A variety of comparing means can be used to compare a target sequence or target motif with the data storage means to identify sequence fragments of the *Borrelia burgdorferi* genome. In the present examples, implementing software which implement the BLAST and BLAZE algorithms, described in Altschul *et al.*, *J. Mol. Biol.* 215: 403-410 (1990), is used to identify open reading frames within the *Borrelia burgdorferi* genome. A skilled artisan can readily recognize that any one of the publicly available homology search programs can be used as the search means for the computer-based systems of the present invention. Of course, suitable proprietary systems that may be known to those of skill also may be employed in this regard.

Figure 1 provides a block diagram of a computer system illustrative of embodiments of this aspect of present invention. The computer system 102 includes a processor 106 connected to a bus 104. Also connected to the bus 104 are a main memory 108 (preferably implemented as random access memory, RAM) and a variety of secondary storage devices 110, such as a hard drive 112 and a removable medium storage device 114. The removable medium storage device

114 may represent, for example, a floppy disk drive, a CD-ROM drive, a magnetic tape drive, etc. A removable storage medium 116 (such as a floppy disk, a compact disk, a magnetic tape, etc.) containing control logic and/or data recorded therein may be inserted into the removable medium storage device 114. The computer system 102 includes appropriate software for reading the control logic and/or the data from the removable medium storage device 114, once it is inserted into the removable medium storage device 114.

A nucleotide sequence of the present invention may be stored in a well known manner in the main memory 108, any of the secondary storage devices 110, and/or a removable storage medium 116. During execution, software for accessing and processing the genomic sequence (such as search tools, comparing tools, etc.) reside in main memory 108, in accordance with the requirements and operating parameters of the operating system, the hardware system and the software program or programs.

BIOCHEMICAL EMBODIMENTS

Other embodiments of the present invention are directed to isolated fragments of the *Borrelia burgdorferi* genome. The fragments of the *Borrelia burgdorferi* genome of the present invention include, but are not limited to fragments which encode peptides, hereinafter open reading frames (ORFs), fragments which modulate the expression of an operably linked ORF, hereinafter expression modulating fragments (EMFs) and fragments which can be used to diagnose the presence of *Borrelia burgdorferi* in a sample, hereinafter diagnostic fragments (DFs).

As used herein, an "isolated nucleic acid molecule" or an "isolated fragment of the *Borrelia burgdorferi* genome" refers to a nucleic acid molecule possessing a specific nucleotide sequence which has been subjected to purification means to reduce, from the composition, the number of compounds which are normally associated with the composition. Particularly, the term refers to the nucleic acid molecules having the sequences set out in SEQ ID NOS: 1-155, to representative fragments thereof as described above including ORF IDs and ORFs, to polynucleotides at least 95%, preferably at least 96%, 97%, 98%, or 99% and especially preferably at least 99.9% identical in sequence thereto, also as set out above.

A variety of purification means can be used to generate the isolated fragments of the present invention. These include, but are not limited to methods which separate constituents of a solution based on charge, solubility, or size.

In one embodiment, *Borrelia burgdorferi* DNA can be enzymatically sheared to produce fragments of 15-20 kb in length. These fragments can then be used to generate a *Borrelia burgdorferi* library by inserting them into lambda clones as described in the Examples below. Primers flanking, for example, an ORF, such as those enumerated in Tables 1-6 can then be generated using nucleotide sequence information provided in SEQ ID NOS: 1-155. Well known and routine techniques of PCR cloning then can be used to isolate the ORF from the lambda DNA library or *Borrelia burgdorferi* genomic DNA. Thus, given the availability of SEQ ID NOS:1-

155, the information in Tables 1-6, and the information that may be obtained readily by analysis of the sequences of SEQ ID NOS:1-155 using methods set out above, those of skill will be enabled by the present disclosure to isolate any ORF-containing or other nucleic acid fragment of the present invention.

5 The isolated nucleic acid molecules of the present invention include, but are not limited to single stranded and double stranded DNA, and single stranded RNA. For purposes of numbering and reference to polynucleotide and polypeptide sequences the entire sequence of each sequence of SEQ ID NOS:1-155 is included with the first nucleotide being position 1.

Therefore, for reference purposes the numbering used in the present invention is that provided in the sequence listing for SEQ ID NOS:1-155.

10 As used herein, an open reading frame (ORF), means a series of nucleotide triplets coding for amino acid residues without any termination codons and is a sequence translatable into protein. Further, unless specified, the term "ORF" for each ORF ID is defined by the termination codon at the 3' end and the 5' most methionine codon, at the 5' end, in frame with said 3' termination codon. Unless specified, the term "ORF" also refers to a particular polypeptide sequence defined by the ORF polynucleotide sequence, wherein the N-terminus is defined by the 5' most methionine codon in frame with the termination codon at the 3' end of the ORF ID and the C-terminus is defined by the last codon before the said 3' termination codon. As used herein, an ORF ID represents a sequence without any internal termination codons flanked by termination codons.

20 Tables 1-6 list ORF IDs in the *Borrelia burgdorferi* genomic contigs of the present invention that were identified as putative coding regions by the GeneMark software using organism-specific second-order Markov probability transition matrices. It will be appreciated that other criteria can be used, in accordance with well known analytical methods, such as those discussed herein, to generate more inclusive, more restrictive, or more selective lists.

25 The *B. burgdorferi* genome consists of one large linear chromosome containing approximately two thirds of its genetic material and multiple extrachromosomal elements (approximately 15) containing the remaining one third of its genetic material. SEQ ID NO:1 (Contig ID 1) is the complete sequence of the large linear *B. burgdorferi* chromosome. SEQ ID NOS:2-155 (Contig ID 2-155 respectively) are fragments (contigs) of the extrachromosomal elements. Tables 1-3 below relate only to SEQ ID NO:1. Tables 4-6 relate to the extrachromosomal elements (SEQ ID NOS:2-155).

35 Table 1 sets out ORF IDs in the *Borrelia burgdorferi* chromosome of the present invention that cover a continuous region of at least 50 bases are 95% or more identical (by BLAST analysis using default parameters) to a nucleotide sequence available through GenBank in July, 1997.

Table 2 sets out ORF IDs in the *Borrelia burgdorferi* chromosome of the present invention that are not in Table 1 and match, with a BLASTP probability score of 0.01 or less, a polypeptide sequence available through GenBank in July, 1997.

Table 3 sets out ORF IDs in the *Borrelia burgdorferi* chromosome of the present invention that do not match significantly, by BLASTP analysis, a polypeptide sequence available through GenBank in July, 1997.

5 Table 4 sets out ORF IDs in the *Borrelia burgdorferi* extrachromosomal element contigs of the present invention that over a continuous region of at least 50 bases are 95% or more identical (by BLAST analysis) to a nucleotide sequence available through GenBank in July, 1997.

10 Table 5 sets out ORF IDs in the *Borrelia burgdorferi* extrachromosomal element contigs of the present invention that are not in Table 1 and match, with a BLASTP probability score of 0.01 or less, a polypeptide sequence available through GenBank in July, 1997.

Table 6 sets out ORF IDs in the *Borrelia burgdorferi* extrachromosomal element contigs of the present invention that do not match significantly, by BLASTP analysis, a polypeptide sequence available through GenBank in July, 1997.

15 In each table, the first and second columns identify the ORF ID by, respectively, contig number and ORF ID number within the contig; the third column indicates the first nucleotide of the ORF ID, counting from the 5' end of the contig strand; and the fourth column indicates the last nucleotide of the ORF ID, counting from the 5' end of the contig strand.

20 In Tables 1, 2, 4 and 5, column five, lists the Reference for the closest matching sequence available through GenBank. These reference numbers are the database accession numbers commonly used by those of skill in the art, who will be familiar with their denominators. Descriptions of the nomenclature are available from the National Center for Biotechnology Information. Column seven provides the BLAST identity score from the comparison of the ORF ID and the homologous gene; and column nine indicates the length in nucleotides of the highest scoring segment pair identified by the BLAST identity analysis.

25 The concepts of percent identity and percent similarity of two polypeptide sequences is well understood in the art. For example, two polypeptides 10 amino acids in length which differ at three amino acid positions (*e.g.*, at positions 1, 3 and 5) are said to have a percent identity of 70%. However, the same two polypeptides would be deemed to have a percent similarity of 80% if, for example at position 5, the amino acids moieties, although not identical, were
30 "similar" (*i.e.*, possessed similar biochemical characteristics). As is known in the art, substitution of one amino acid for a "similar" amino acid is a conservative substitution. Generally, proteins are highly tolerant of conservative substitutions. Many programs for analysis of nucleotide or amino acid sequence similarity, such as fasta and BLAST specifically list percent identity of a matching region as an output parameter. Thus, for instance, Tables 1, 2, 4 and 5
35 herein enumerate the percent identity and similarity of the highest scoring segment pair in each ORF and its listed relative. Further details concerning the algorithms and criteria used for homology searches are provided below and are described in the pertinent literature highlighted by the citations provided below.

It will be appreciated that other criteria can be used to generate more inclusive and more exclusive listings of the types set out in the tables. As those of skill will appreciate, narrow and broad searches both are useful. Thus, a skilled artisan can readily identify ORFs in contigs of the *Borrelia burgdorferi* genome other than those listed in Tables 1-6, such as ORFs which are overlapping or encoded by the opposite strand of an identified ORF in addition to those ascertainable using the computer-based systems of the present invention.

As used herein, an "expression modulating fragment," EMF, means a series of nucleotide molecules which modulates the expression of an operably linked ORF or EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

EMF sequences can be identified within the contigs of the *Borrelia burgdorferi* genome by their proximity to the ORFs provided in Tables 1-6. An intergenic segment, or a fragment of the intergenic segment, from about 10 to 200 nucleotides in length, taken from any one of the ORFs of Tables 1-6 will modulate the expression of an operably linked ORF in a fashion similar to that found with the naturally linked ORF sequence. As used herein, an "intergenic segment" refers to fragments of the *Borrelia burgdorferi* genome which are between two ORF(s) herein described. EMFs also can be identified using known EMFs as a target sequence or target motif in the computer-based systems of the present invention. Further, the two methods can be combined and used together.

The presence and activity of an EMF can be confirmed using an EMF trap vector. An EMF trap vector contains a cloning site linked to a marker sequence. A marker sequence encodes an identifiable phenotype, such as antibiotic resistance or a complementing nutrition auxotrophic factor, which can be identified or assayed when the EMF trap vector is placed within an appropriate host under appropriate conditions. As described above, a EMF will modulate the expression of an operably linked marker sequence. A more detailed discussion of various marker sequences is provided below. A sequence which is suspected as being an EMF is cloned in all three reading frames in one or more restriction sites upstream from the marker sequence in the EMF trap vector. The vector is then transformed into an appropriate host using known procedures and the phenotype of the transformed host is examined under appropriate conditions. As described above, an EMF will modulate the expression of an operably linked marker sequence.

As used herein, a "diagnostic fragment," DF, means a series of nucleotide molecules which selectively hybridize to *Borrelia burgdorferi* sequences. DFs can be readily identified by identifying unique sequences within contigs of the *Borrelia burgdorferi* genome, such as by using well-known computer analysis software, and by generating and testing probes or

amplification primers consisting of the DF sequence in an appropriate diagnostic format which determines amplification or hybridization selectivity.

The sequences falling within the scope of the present invention are not limited to the specific sequences herein described, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequences provided in SEQ ID NOS:1-155, ORF IDs and ORFs within, a representative fragment thereof, or a nucleotide sequence at least 99% and preferably 99.9% identical to SEQ ID NOS: 1-155, ORF IDs and ORFs within, with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another which encodes the same amino acid is expressly contemplated.

Any specific sequence disclosed herein can be readily screened for errors by resequencing a particular fragment, such as an ORF, in both directions (*i.e.*, sequence both strands).

Alternatively, error screening can be performed by sequencing corresponding polynucleotides of *Borrelia burgdorferi* origin isolated by using part or all of the fragments in question as a probe or primer.

Each of the ORF IDs and ORFs of the *Borrelia burgdorferi* genome disclosed in Tables 1-6, and the EMFs found 5' to the ORF IDs, can be used as polynucleotide reagents in numerous ways. For example, the sequences can be used as diagnostic probes or diagnostic amplification primers to detect the presence of a specific microbe in a sample, particularly *Borrelia burgdorferi*. Especially preferred in this regard are ORF IDs and ORFs such as those of Tables 3 and 6, which do not match previously characterized sequences from other organisms and thus are most likely to be highly selective for *Borrelia burgdorferi*. Also particularly preferred are ORF IDs and ORFs that can be used to distinguish between strains of *Borrelia burgdorferi*, particularly those that distinguish medically important strain, such as drug-resistant strains.

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA. Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Information from the sequences of the present invention can be used to design antisense and triple helix-forming oligonucleotides. Polynucleotides suitable for use in these methods are usually 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription, for triple-helix formation, or to the mRNA itself, for antisense inhibition. Both techniques have been demonstrated to be effective in model systems, and the requisite techniques are well known and involve routine procedures. Triple helix techniques are discussed in, for example, Lee *et al.*, *Nucl. Acids Res.* 6:3073 (1979); Cooney *et al.*, *Science* 241:456 (1988); and Dervan *et al.*, *Science* 251:1360 (1991). Antisense techniques in general are discussed in, for instance, Okano,

J. Neurochem. 56:560 (1991) and *Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression*, CRC Press, Boca Raton, FL (1988)).

The present invention further provides recombinant constructs comprising one or more fragments of the *Borrelia burgdorferi* genomic fragments and contigs of the present invention.

5 Certain preferred recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a fragment of the *Borrelia burgdorferi* genome has been inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORF IDs or ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF ID or ORF. For vectors
10 comprising the EMFs of the present invention, the vector may further comprise a marker sequence or heterologous ORF ID or ORF operably linked to the EMF.

Large numbers of suitable vectors and promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Useful bacterial vectors include
15 phagescript, PsiX174, pBluescript SK, pBS KS, pNH8a, pNH16a, pNH18a, pNH46a (available from Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (available from Pharmacia); pQE vectors (available from Promega). Useful eukaryotic vectors include pWLneo, pSV2cat, pOG44, pXT1, pSG (available from Stratagene) pSVK3, pBPV, pMSG, pSVL (available from Pharmacia).

20 Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein- I. Selection of
25 the appropriate vector and promoter is well within the level of ordinary skill in the art.

The present invention further provides host cells containing any one of the isolated fragments of the *Borrelia burgdorferi* genomic fragments and contigs of the present invention, wherein the fragment has been introduced into the host cell using known methods. The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such
30 as a yeast cell, or a procaryotic cell, such as a bacterial cell.

A polynucleotide of the present invention, such as a recombinant construct comprising an ORF of the present invention, may be introduced into the host by a variety of well established techniques that are standard in the art, such as calcium phosphate transfection, DEAE, dextran mediated transfection and electroporation, which are described in, for instance, Davis, L. *et al.*,
35 BASIC METHODS IN MOLECULAR BIOLOGY (1986).

A host cell containing one of the fragments of the *Borrelia burgdorferi* genomic fragments and contigs of the present invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (*e.g.*, an ORF) by nucleotide sequence but, due to the degeneracy of the Genetic Code, encode an identical polypeptide sequence.

Preferred nucleic acid fragments of the present invention are the ORF IDs depicted in Tables 2, 3, 5 and 6, and ORFs within, which encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. This is particularly useful in producing small peptides and fragments of larger polypeptides. Such short fragments as may be obtained most readily by synthesis are useful, for example, in generating antibodies against the native polypeptide, as discussed further below.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily employ well-known methods for isolating polypeptides and proteins to isolate and purify polypeptides or proteins of the present invention produced naturally by a bacterial strain, or by other methods. Methods for isolation and purification that can be employed in this regard include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography.

The polypeptides and proteins of the present invention also can be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. Those skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of the *B. burgdorferi* polypeptide can be substantially purified by the one-step method described by Smith et al. (1988) Gene 67:31-40. Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies directed against the polypeptides of the invention in methods which are well known in the art of protein purification.

The invention further provides for isolated *B. burgdorferi* polypeptides comprising an amino acid sequence selected from the group including: (a) the amino acid sequence of a full-length *B. burgdorferi* polypeptide having the complete amino acid sequence from the first methionine codon to the termination codon of each sequence listed in SEQ ID NOS:1-155, wherein said termination codon is at the end of each SEQ ID NO: and said first methionine is the

first methionine in frame with said termination codon; and (b) the amino acid sequence of a full-length *B. burgdorferi* polypeptide having the complete amino acid sequence in (a) excepting the N-terminal methionine.

The polypeptides of the present invention also include polypeptides having an amino acid sequence at least 80% identical, more preferably at least 90% identical, and still more preferably 95%, 96%, 97%, 98% or 99% identical to those described in (a) and (b) above.

The present invention is further directed to polynucleotides encoding portions or fragments of the amino acid sequences described herein as well as to portions or fragments of the isolated amino acid sequences described herein. Fragments include portions of the amino acid sequences described herein at least 5 contiguous amino acid in length and selected from any two integers, one of which representing an N-terminal position and another representing a C-terminal position. The initiation codon of the ORFs of the present invention is position 1. The initiation codon (position 1) for purposes of the present invention is the first methionine codon of each ORF ID which is in frame with the termination codon at the end of each said sequence. Every combination of a N-terminal and C-terminal position that a fragment at least 5 contiguous amino acid residues in length could occupy, on any given ORF is included in the invention, i.e., from initiation codon up to the termination codon. "At least" means a fragment may be 5 contiguous amino acid residues in length or any integer between 5 and the number of residues in an ORF, minus 1. Therefore, included in the invention are contiguous fragments specified by any N-terminal and C-terminal positions of amino acid sequence set forth in SEQ ID NOS:1-155 or Tables 1-6 wherein the contiguous fragment is any integer between 5 and the number of residues in an ORF minus 1.

Further, the invention includes polypeptides comprising fragments specified by size, in amino acid residues, rather than by N-terminal and C-terminal positions. The invention includes any fragment size, in contiguous amino acid residues, selected from integers between 5 and the number of residues in an ORF, minus 1. Preferred sizes of contiguous polypeptide fragments include about 5 amino acid residues, about 10 amino acid residues, about 20 amino acid residues, about 30 amino acid residues, about 40 amino acid residues, about 50 amino acid residues, about 100 amino acid residues, about 200 amino acid residues, about 300 amino acid residues, and about 400 amino acid residues. The preferred sizes are, of course, meant to exemplify, not limit, the present invention as all size fragments representing any integer between 5 and the number of residues in a full length sequence minus 1 are included in the invention. The present invention also provides for the exclusion of any fragments specified by N-terminal and C-terminal positions or by size in amino acid residues as described above. Any number of fragments specified by N-terminal and C-terminal positions or by size in amino acid residues as described above may be excluded.

The above fragments need not be active since they would be useful, for example, in immunoassays, in epitope mapping, epitope tagging, to generate antibodies to a particular portion of the protein, as vaccines, and as molecular weight markers.

Further polypeptides of the present invention include polypeptides which have at least 90% similarity, more preferably at least 95% similarity, and still more preferably at least 96%, 97%, 98% or 99% similarity to those described above.

5 A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of a *B. burgdorferi* polypeptide having an amino acid sequence which contains at least one conservative amino acid substitution, but not more than 50 conservative amino acid substitutions, not more than 40 conservative amino acid substitutions, not more than 30 conservative amino acid substitutions, and not more than 20 conservative amino acid substitutions. Also provided are polypeptides which comprise the amino acid sequence of a *B.*
10 *burgdorferi* polypeptide, having at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 conservative amino acid substitutions.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide
15 sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the
20 reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to the ORF amino acid sequences encoded by the sequences of SEQ
25 ID NOS:1-155, as described hererin, can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al., (1990) Comp. App. Biosci. 6:237-245. In a sequence alignment the query and
30 subject sequences are both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is
35 shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, the results, in percent identity, must be manually corrected. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject

sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is
5 determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the
10 purposes of manually adjusting the percent identity score. That is, only query amino acid residues outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not match/align with the first 10
15 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence.
20 This time the deletions are internal so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected. No other manual
25 corrections are to made for the purposes of the present invention.

The above polypeptide sequences are included irrespective of whether they have their normal biological activity. This is because even where a particular polypeptide molecule does not have biological activity, one of skill in the art would still know how to use the polypeptide, for instance, as a vaccine or to generate antibodies. Other uses of the polypeptides of the present
30 invention that do not have *B. burgdorferi* activity include, *inter alia*, as epitope tags, in epitope mapping, and as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods known to those of skill in the art.

As described below, the polypeptides of the present invention can also be used to raise polyclonal and monoclonal antibodies, which are useful in assays for detecting *B. burgdorferi* protein
35 expression or as agonists and antagonists capable of enhancing or inhibiting *B. burgdorferi* protein function. Further, such polypeptides can be used in the yeast two-hybrid system to "capture" *B. burgdorferi* protein binding proteins which are also candidate agonists and antagonists according to the present invention. See, e.g., Fields et al. (1989) Nature 340:245-246.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, CV-1 cell, COS cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level.

"Recombinant," as used herein, means that a polypeptide or protein is derived from recombinant (*e.g.*, microbial or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (*e.g.*, yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, *e.g.*, *E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern different from that expressed in mammalian cells.

"Nucleotide sequence" refers to a heteropolymer of deoxyribonucleotides. Generally, DNA segments encoding the polypeptides and proteins provided by this invention are assembled from fragments of the *Borrelia burgdorferi* genome and short oligonucleotide linkers, or from a series of oligonucleotides, to provide a synthetic gene which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon.

Recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. The expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic regulatory elements necessary for gene expression in the host, including elements required to initiate and maintain transcription at a level sufficient for suitable expression of the desired polypeptide, including, for example, promoters and, where necessary, an enhancer and a polyadenylation signal; (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate signals to initiate translation at the beginning of the desired coding region and terminate translation at its end. Structural units intended for use in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an N-terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

"Recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extra chromosomally. The cells can be prokaryotic or eukaryotic. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to

produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference in its entirety.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), alpha-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and, when desirable, provide amplification within the host.

Suitable prokaryotic hosts for transformation include strains of *E. coli*, *B. subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas* and *Streptomyces*. Others may, also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (available from Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (available from Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter, where it is inducible, is derepressed or induced by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period to provide for expression of the induced gene product. Thereafter cells are typically harvested, generally by centrifugation, disrupted to release expressed protein, generally by physical or chemical means, and the resulting crude extract is retained for further purification.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney

fibroblasts, described in Gluzman, *Cell* 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines.

Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

Recombinant polypeptides and proteins produced in bacterial culture is usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The present invention further includes isolated polypeptides, proteins and nucleic acid molecules which are substantially equivalent to those herein described. As used herein, substantially equivalent can refer both to nucleic acid and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between reference and subject sequences. Particularly preferred in this regard are conservative substitutions, known to those of skill in the art. For purposes of the present invention, sequences having equivalent biological activity, and equivalent expression characteristics are considered substantially equivalent. For purposes of determining equivalence, truncation of the mature sequence (e.g., removal of leader sequence(s)) should be disregarded.

The invention further provides methods of obtaining homologs from other strains of *Borrelia burgdorferi*, of the fragments of the *Borrelia burgdorferi* genome of the present invention and homologs of the proteins encoded by the ORFs of the present invention. As used herein, a sequence or protein of *Borrelia burgdorferi* is defined as a homolog of a fragment of the *Borrelia burgdorferi* fragments or contigs or a protein encoded by one of the ORFs of the present invention, if it shares significant homology to one of the fragments of the *Borrelia burgdorferi* genome of the present invention or a protein encoded by one of the ORFs of the present invention. Specifically, by using the sequence disclosed herein as a probe or as primers, and techniques such as PCR cloning and colony/plaque hybridization, one skilled in the art can obtain homologs.

As used herein, two nucleic acid molecules or proteins are said to "share significant homology" if the two contain regions which possess greater than 85% sequence (amino acid or nucleic acid) homology. Preferred homologs in this regard are those with more than 90% homology. Especially preferred are those with 95% or more homology. Among especially

preferred homologs those with 96, 97%, 98%, 99% or more homology are particularly preferred. The most preferred homologs among these are those with 99.9% homology or more. It will be understood that, among measures of homology, identity is particularly preferred in this regard.

5 Region specific primers or probes derived from the nucleotide sequence provided in SEQ ID NOS: 1-155 or from a nucleotide sequence at least 95%, particularly at least 96%, 97%, 98% or 99%, especially at least 99.5% identical to a sequence of SEQ ID NOS: 1-155 can be used to prime DNA synthesis and PCR amplification, as well as to identify colonies containing cloned DNA encoding a homolog. Methods suitable to this aspect of the present invention are well
10 known and have been described in great detail in many publications such as, for example, Innis *et al.*, *PCR Protocols*, Academic Press, San Diego, CA (1990)).

 When using primers derived from SEQ ID NOS: 1-155 or from a nucleotide sequence having an aforementioned identity to a sequence of SEQ ID NOS:1-155, one skilled in the art will recognize that by employing high stringency conditions (*e.g.*, annealing at 50-60°C in 6X SSPC
15 and 50% formamide, and washing at 50- 65°C in 0.5X SSPC) only sequences which are greater than 75% homologous to the primer will be amplified. By employing lower stringency conditions (*e.g.*, hybridizing at 35-37°C in 5X SSPC and 40-45% formamide, and washing at 42°C in 0.5X SSPC), sequences which are greater than 40-50% homologous to the primer will also be amplified.

20 When using DNA probes derived from SEQ ID NOS:1-155, or from a nucleotide sequence having an aforementioned identity to a sequence of SEQ ID NOS: 1-155 , for colony/plaque hybridization, one skilled in the art will recognize that by employing high stringency conditions (*e.g.*, hybridizing at 50- 65°C in 5X SSPC and 50% formamide, and washing at 50- 65°C in 0.5X SSPC), sequences having regions which are greater than 90%
25 homologous to the probe can be obtained, and that by employing lower stringency conditions (*e.g.*, hybridizing at 35-37°C in 5X SSPC and 40-45% formamide, and washing at 42°C in 0.5X SSPC), sequences having regions which are greater than 35-45% homologous to the probe will be obtained.

 Any organism can be used as the source for homologs of the present invention so long as
30 the organism naturally expresses such a protein or contains genes encoding the same. The most preferred organism for isolating homologs are bacteria which are closely related to *Borrelia burgdorferi*.

ILLUSTRATIVE USES OF COMPOSITIONS OF THE INVENTION

35 Each ORF of the ORF IDs provided in Tables 1, 2, 4 and 5 is identified with a function by homology to a known gene or polypeptide. As a result, one skilled in the art can use the polypeptides of the present invention for commercial, therapeutic and industrial purposes consistent with the type of putative identification of the polypeptide. Such identifications permit one skilled in the art to use the *Borrelia burgdorferi* ORFs in a manner similar to the known type

of sequences for which the identification is made; for example, to ferment a particular sugar source or to produce a particular metabolite. A variety of reviews illustrative of this aspect of the invention are available, including the following reviews on the industrial use of enzymes, for example, BIOCHEMICAL ENGINEERING AND BIOTECHNOLOGY HANDBOOK, 2nd Ed., MacMillan Publications, Ltd. NY (1991) and BIOCATALYSTS IN ORGANIC SYNTHESIS, Tramper *et al.*, Eds., Elsevier Science Publishers, Amsterdam, The Netherlands (1985). A variety of exemplary uses that illustrate this and similar aspects of the present invention are discussed below.

1. Biosynthetic Enzymes

Open reading frames encoding proteins involved in mediating the catalytic reactions involved in intermediary and macromolecular metabolism, the biosynthesis of small molecules, cellular processes and other functions includes enzymes involved in the degradation of the intermediary products of metabolism, enzymes involved in central intermediary metabolism, enzymes involved in respiration, both aerobic and anaerobic, enzymes involved in fermentation, enzymes involved in ATP proton motor force conversion, enzymes involved in broad regulatory function, enzymes involved in amino acid synthesis, enzymes involved in nucleotide synthesis, enzymes involved in cofactor and vitamin synthesis, can be used for industrial biosynthesis.

The various metabolic pathways present in *Borrelia burgdorferi* can be identified based on absolute nutritional requirements as well as by examining the various enzymes identified in Table 1-6 and SEQ ID NOS:1-155.

Of particular interest are polypeptides involved in the degradation of intermediary metabolites as well as non-macromolecular metabolism. Such enzymes include amylases, glucose oxidases, and catalase.

Proteolytic enzymes are another class of commercially important enzymes. Proteolytic enzymes find use in a number of industrial processes including the processing of flax and other vegetable fibers, in the extraction, clarification and depectinization of fruit juices, in the extraction of vegetables' oil and in the maceration of fruits and vegetables to give unicellular fruits. A detailed review of the proteolytic enzymes used in the food industry is provided in Rombouts *et al.*, *Symbiosis* 21:79 (1986) and Voragen *et al.* in *Biocatalysts In Agricultural Biotechnology*, Whitaker *et al.*, Eds., *American Chemical Society Symposium Series* 389:93 (1989).

The metabolism of sugars is an important aspect of the primary metabolism of *Borrelia burgdorferi*. Enzymes involved in the degradation of sugars, such as, particularly, glucose, galactose, fructose and xylose, can be used in industrial fermentation. Some of the important sugar transforming enzymes, from a commercial viewpoint, include sugar isomerases such as glucose isomerase. Other metabolic enzymes have found commercial use such as glucose oxidases which produces ketogulonic acid (KGA). KGA is an intermediate in the commercial production of ascorbic acid using the Reichstein's procedure, as described in Krueger *et al.*, *Biotechnology* 6(A), Rhine *et al.*, Eds., Verlag Press, Weinheim, Germany (1984).

Glucose oxidase (GOD) is commercially available and has been used in purified form as well as in an immobilized form for the deoxygenation of beer. See, for instance, Hartmeir *et al.*, *Biotechnology Letters* 1:21 (1979). The most important application of GOD is the industrial scale fermentation of gluconic acid. Market for gluconic acids which are used in the detergent, textile, leather, photographic, pharmaceutical, food, feed and concrete industry, as described, for example, in Bigelis *et al.*, beginning on page 357 in GENE MANIPULATIONS AND FUNGI; Bennett *et al.*, Eds., Academic Press, New York (1985). In addition to industrial applications, GOD has found applications in medicine for quantitative determination of glucose in body fluids recently in biotechnology for analyzing syrups from starch and cellulose hydrosylates. This application is described in Owusu *et al.*, *Biochem. et Biophysica. Acta.* 872:83 (1986), for instance.

The main sweetener used in the world today is sugar which comes from sugar beets and sugar cane. In the field of industrial enzymes, the glucose isomerase process shows the largest expansion in the market today. Initially, soluble enzymes were used and later immobilized enzymes were developed (Krueger *et al.*, *Biotechnology, The Textbook of Industrial Microbiology*, Sinauer Associated Incorporated, Sunderland, Massachusetts (1990)). Today, the use of glucose- produced high fructose syrups is by far the largest industrial business using immobilized enzymes. A review of the industrial use of these enzymes is provided by Jorgensen, *Starch* 40:307 (1988).

Proteinases, such as alkaline serine proteinases, are used as detergent additives and thus represent one of the largest volumes of microbial enzymes used in the industrial sector. Because of their industrial importance, there is a large body of published and unpublished information regarding the use of these enzymes in industrial processes. (See Faultman *et al.*, *Acid Proteases Structure Function and Biology*, Tang, J., ed., Plenum Press, New York (1977) and Godfrey *et al.*, *Industrial Enzymes*, MacMillan Publishers, Surrey, UK (1983) and Hepner *et al.*, *Report Industrial Enzymes by 1990*, Hel Hepner & Associates, London (1986)).

Another class of commercially usable proteins of the present invention are the microbial lipases, described by, for instance, Macrae *et al.*, *Philosophical Transactions of the Chiral Society of London* 310:227 (1985) and Poserke, *Journal of the American Oil Chemist Society* 61:1758 (1984). A major use of lipases is in the fat and oil industry for the production of neutral glycerides using lipase catalyzed inter-esterification of readily available triglycerides. Application of lipases include the use as a detergent additive to facilitate the removal of fats from fabrics in the course of the washing procedures.

The use of enzymes, and in particular microbial enzymes, as catalyst for key steps in the synthesis of complex organic molecules is gaining popularity at a great rate. One area of great interest is the preparation of chiral intermediates. Preparation of chiral intermediates is of interest to a wide range of synthetic chemists particularly those scientists involved with the preparation of new pharmaceuticals, agrochemicals, fragrances and flavors. (See Davies *et al.*, *Recent Advances in the Generation of Chiral Intermediates Using Enzymes*, CRC Press, Boca Raton,

Florida (1990)). The following reactions catalyzed by enzymes are of interest to organic chemists: hydrolysis of carboxylic acid esters, phosphate esters, amides and nitriles, esterification reactions, trans-esterification reactions, synthesis of amides, reduction of alkanones and oxoalkanates, oxidation of alcohols to carbonyl compounds, oxidation of sulfides to sulfoxides, and carbon bond forming reactions such as the aldol reaction.

When considering the use of an enzyme encoded by one of the ORFs of the present invention for biotransformation and organic synthesis it is sometimes necessary to consider the respective advantages and disadvantages of using a microorganism as opposed to an isolated enzyme. Pros and cons of using a whole cell system on the one hand or an isolated partially purified enzyme on the other hand, has been described in detail by Bud *et al.*, Chemistry in Britain (1987), p. 127.

Amino transferases, enzymes involved in the biosynthesis and metabolism of amino acids, are useful in the catalytic production of amino acids. The advantages of using microbial based enzyme systems is that the amino transferase enzymes catalyze the stereo-selective synthesis of only L-amino acids and generally possess uniformly high catalytic rates. A description of the use of amino transferases for amino acid production is provided by Roselle-David, *Methods of Enzymology* 136:479 (1987).

Another category of useful proteins encoded by the ORFs of the present invention include enzymes involved in nucleic acid synthesis, repair, and recombination.

2. Generation of Antibodies

As described here, the proteins of the present invention, as well as homologs thereof, can be used in a variety of procedures and methods known in the art which are currently applied to other proteins. The proteins of the present invention can further be used to generate an antibody which selectively binds the protein.

B. burgdorferi protein-specific antibodies for use in the present invention can be raised against the intact *B. burgdorferi* protein or an antigenic polypeptide fragment thereof, which may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier.

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules, single chain whole antibodies, and antibody fragments. Antibody fragments of the present invention include Fab and F(ab')₂ and other fragments including single-chain Fvs (scFv) and disulfide-linked Fvs (sdFv). Also included in the present invention are chimeric and humanized monoclonal antibodies and polyclonal antibodies specific for the polypeptides of the present invention. The antibodies of the present invention may be prepared by any of a variety of methods. For example, cells expressing a polypeptide of the present invention or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies. For example, a preparation of *B. burgdorferi* polypeptide or fragment thereof is prepared and purified to render it substantially free

of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In a preferred method, the antibodies of the present invention are monoclonal antibodies or binding fragments thereof. Such monoclonal antibodies can be prepared using hybridoma technology. See, e.g., Harlow et al., ANTIBODIES: A LABORATORY MANUAL, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling, et al., in: MONOCLONAL ANTIBODIES AND T-CELL HYBRIDOMAS 563-681 (Elsevier, N.Y., 1981). Fab and F(ab')₂ fragments may be produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Alternatively, *B. burgdorferi* polypeptide-binding fragments, chimeric, and humanized antibodies can be produced through the application of recombinant DNA technology or through synthetic chemistry using methods known in the art.

Alternatively, additional antibodies capable of binding to the polypeptide antigen of the present invention may be produced in a two-step procedure through the use of anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and that, therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, *B. burgdorferi* polypeptide-specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the *B. burgdorferi* polypeptide-specific antibody can be blocked by the *B. burgdorferi* polypeptide antigen. Such antibodies comprise anti-idiotypic antibodies to the *B. burgdorferi* polypeptide-specific antibody and can be used to immunize an animal to induce formation of further *B. burgdorferi* polypeptide-specific antibodies.

Antibodies and fragments thereof of the present invention may be described by the portion of a polypeptide of the present invention recognized or specifically bound by the antibody. Antibody binding fragments of a polypeptide of the present invention may be described or specified in the same manner as for polypeptide fragments discussed above, i.e., by N-terminal and C-terminal positions or by size in contiguous amino acid residues. Any number of antibody binding fragments, of a polypeptide of the present invention, specified by N-terminal and C-terminal positions or by size in amino acid residues, as described above, may also be excluded from the present invention. Therefore, the present invention includes antibodies that specifically bind a particular discribed fragment of a polypeptide of the present invention and allows for the exclusion of the same.

Antibodies and fragments thereof of the present invention may also be described or specified in terms of their cross-reactivity. Antibodies and fragments that do not bind polypeptides of any other species of *Borrelia* other than *B. burgdorferi* are included in the present invention. Likewise, antibodies and fragments that bind only species of *Borrelia*, i.e. antibodies and fragments that do not bind bacteria from any genus other than *Borrelia*, are included in the present invention.

3. Epitope-Bearing Portions

In another aspect, the invention provides peptides and polypeptides comprising epitope-bearing portions of the *B. burgdorferi* polypeptides of the present invention. These epitopes are immunogenic or antigenic epitopes of the polypeptides of the present invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein or polypeptide is the immunogen. These immunogenic epitopes are believed to be confined to a few loci on the molecule. On the other hand, a region of a protein molecule to which an antibody can bind is defined as an "antigenic determinant" or "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes. See, e.g., Geysen, et al. (1983) Proc. Natl. Acad. Sci. USA 81:3998- 4002. Amino acid residues comprising antigenic epitopes may be determined by algorithms such as the Jameson-Wolf analysis or similar algorithms or by *in vivo* testing for an antigenic response using the methods described herein or those known in the art.

As to the selection of peptides or polypeptides bearing an antigenic epitope (*i.e.*, that contain a region of a protein molecule to which an antibody can bind), it is well known in that art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein. See, e.g., Sutcliffe, et al., (1983) Science 219:660-666. Peptides capable of eliciting protein-reactive sera are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are confined neither to immunodominant regions of intact proteins (*i.e.*, immunogenic epitopes) nor to the amino or carboxyl terminals. Peptides that are extremely hydrophobic and those of six or fewer residues generally are ineffective at inducing antibodies that bind to the mimicked protein; longer, peptides, especially those containing proline residues, usually are effective. See, Sutcliffe, et al., *supra*, p. 661. For instance, 18 of 20 peptides designed according to these guidelines, containing 8-39 residues covering 75% of the sequence of the influenza virus hemagglutinin HA1 polypeptide chain, induced antibodies that reacted with the HA1 protein or intact virus; and 12/12 peptides from the MuLV polymerase and 18/18 from the rabies glycoprotein induced antibodies that precipitated the respective proteins.

Antigenic epitope-bearing peptides and polypeptides of the invention are therefore useful to raise antibodies, including monoclonal antibodies, that bind specifically to a polypeptide of the invention. Thus, a high proportion of hybridomas obtained by fusion of spleen cells from donors immunized with an antigen epitope-bearing peptide generally secrete antibody reactive with the native protein. See Sutcliffe, et al., *supra*, p. 663. The antibodies raised by antigenic epitope-bearing peptides or polypeptides are useful to detect the mimicked protein, and antibodies to different peptides may be used for tracking the fate of various regions of a protein precursor which undergoes post-translational processing. The peptides and anti-peptide antibodies may be used in a variety of qualitative or quantitative assays for the mimicked protein, for instance in competition assays since it has been shown that even short peptides (*e.g.*, about 9 amino acids)

can bind and displace the larger peptides in immunoprecipitation assays. *See, e.g.,* Wilson, et al., (1984) Cell 37:767-778. The anti-peptide antibodies of the invention also are useful for purification of the mimicked protein, for instance, by adsorption chromatography using methods known in the art.

5 Antigenic epitope-bearing peptides and polypeptides of the invention designed according to the above guidelines preferably contain a sequence of at least seven, more preferably at least nine and most preferably between about 10 to about 50 amino acids (i.e. any integer between 7 and 50) contained within the amino acid sequence of a polypeptide of the invention. However, peptides or polypeptides comprising a larger portion of an amino acid sequence of a polypeptide
10 of the invention, containing about 50 to about 100 amino acids, or any length up to and including the entire amino acid sequence of a polypeptide of the invention, also are considered epitope-bearing peptides or polypeptides of the invention and also are useful for inducing antibodies that react with the mimicked protein. Preferably, the amino acid sequence of the epitope-bearing peptide is selected to provide substantial solubility in aqueous solvents (i.e., the
15 sequence includes relatively hydrophilic residues and highly hydrophobic sequences are preferably avoided); and sequences containing proline residues are particularly preferred.

The epitope-bearing peptides and polypeptides of the present invention may be produced by any conventional means for making peptides or polypeptides including recombinant means using nucleic acid molecules of the invention. For instance, an epitope-bearing amino acid
20 sequence of the present invention may be fused to a larger polypeptide which acts as a carrier during recombinant production and purification, as well as during immunization to produce anti-peptide antibodies. Epitope-bearing peptides also may be synthesized using known methods of chemical synthesis. For instance, Houghten has described a simple method for synthesis of large numbers of peptides, such as 10-20 mg of 248 different 13 residue peptides representing
25 single amino acid variants of a segment of the HA1 polypeptide which were prepared and characterized (by ELISA-type binding studies) in less than four weeks (Houghten, R. A. Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985)). This "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Patent No. 4,631,211 to Houghten and coworkers (1986). In this procedure the individual resins for the solid-phase synthesis of various peptides
30 are contained in separate solvent-permeable packets, enabling the optimal use of the many identical repetitive steps involved in solid-phase methods. A completely manual procedure allows 500-1000 or more syntheses to be conducted simultaneously (Houghten et al. (1985) Proc. Natl. Acad. Sci. 82:5131-5135 at 5134.

Epitope-bearing peptides and polypeptides of the invention are used to induce antibodies
35 according to methods well known in the art. *See, e.g.,* Sutcliffe, et al., *supra*; Wilson, et al., *supra*; and Bittle, et al. (1985) J. Gen. Virol. 66:2347-2354. Generally, animals may be immunized with free peptide; however, anti-peptide antibody titer may be boosted by coupling of the peptide to a macromolecular carrier, such as keyhole limpet hemacyanin (KLH) or tetanus toxoid. For instance, peptides containing cysteine may be coupled to carrier using a linker such

as m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carrier using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice are immunized with either free or carrier-coupled peptides, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 µg peptide or carrier protein and Freund's adjuvant. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of anti-peptide antibody which can be detected, for example, by ELISA assay using free peptide adsorbed to a solid surface. The titer of anti-peptide antibodies in serum from an immunized animal may be increased by selection of anti-peptide antibodies, for instance, by adsorption to the peptide on a solid support and elution of the selected antibodies according to methods well known in the art.

Immunogenic epitope-bearing peptides of the invention, *i.e.*, those parts of a protein that elicit an antibody response when the whole protein is the immunogen, are identified according to methods known in the art. For instance, Geysen, *et al.*, *supra*, discloses a procedure for rapid concurrent synthesis on solid supports of hundreds of peptides of sufficient purity to react in an ELISA. Interaction of synthesized peptides with antibodies is then easily detected without removing them from the support. In this manner a peptide bearing an immunogenic epitope of a desired protein may be identified routinely by one of ordinary skill in the art. For instance, the immunologically important epitope in the coat protein of foot-and-mouth disease virus was located by Geysen *et al. supra* with a resolution of seven amino acids by synthesis of an overlapping set of all 208 possible hexapeptides covering the entire 213 amino acid sequence of the protein. Then, a complete replacement set of peptides in which all 20 amino acids were substituted in turn at every position within the epitope were synthesized, and the particular amino acids conferring specificity for the reaction with antibody were determined. Thus, peptide analogs of the epitope-bearing peptides of the invention can be made routinely by this method. U.S. Patent No. 4,708,781 to Geysen (1987) further describes this method of identifying a peptide bearing an immunogenic epitope of a desired protein.

Further still, U.S. Patent No. 5,194,392, to Geysen (1990), describes a general method of detecting or determining the sequence of monomers (amino acids or other compounds) which is a topological equivalent of the epitope (*i.e.*, a "mimotope") which is complementary to a particular paratope (antigen binding site) of an antibody of interest. More generally, U.S. Patent No. 4,433,092, also to Geysen (1989), describes a method of detecting or determining a sequence of monomers which is a topographical equivalent of a ligand which is complementary to the ligand binding site of a particular receptor of interest. Similarly, U.S. Patent No. 5,480,971 to Houghten, R. A. *et al.* (1996) discloses linear C₁-C₇-alkyl peralkylated oligopeptides and sets and libraries of such peptides, as well as methods for using such oligopeptide sets and libraries for determining the sequence of a peralkylated oligopeptide that preferentially binds to an acceptor molecule of interest. Thus, non-peptide analogs of the epitope-bearing peptides of the invention also can be made routinely by these methods. The entire disclosure of each document cited in this section on "Polypeptides and Fragments" is

hereby incorporated herein by reference.

As one of skill in the art will appreciate, the polypeptides of the present invention and the epitope-bearing fragments thereof described above can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life *in vivo*. This has been shown, *e.g.*, for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EPA 0,394,827; Trauneker et al. (1988) *Nature* 331:84-86. Fusion proteins that have a disulfide-linked dimeric structure due to the IgG part can also be more efficient in binding and neutralizing other molecules than a monomeric *B. burgdorferi* polypeptide or fragment thereof alone. *See* Fountoulakis et al. (1995) *J. Biochem.* 270:3958-3964. Nucleic acids encoding the above epitopes of *B. burgdorferi* polypeptides can also be recombined with a gene of interest as an epitope tag to aid in detection and purification of the expressed polypeptide.

4. Diagnostic Assays and Kits

The present invention further relates to methods for assaying *Borrelia* infection in an animal by detecting the expression of genes encoding *Borrelia* polypeptides of the present invention. The methods comprise analyzing tissue or body fluid from the animal for *Borrelia*-specific antibodies, nucleic acids, or proteins. Analysis of nucleic acid specific to *Borrelia* is assayed by PCR or hybridization techniques using nucleic acid sequences of the present invention as either hybridization probes or primers. *See, e.g.*, Sambrook et al. *Molecular cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 2nd ed., 1989, page 54 reference); Ereemeeva et al. (1994) *J. Clin. Microbiol.* 32:803-810 (describing differentiation among spotted fever group *Rickettsiae* species by analysis of restriction fragment length polymorphism of PCR-amplified DNA) and Chen et al. 1994 *J. Clin. Microbiol.* 32:589-595 (detecting *B. burgdorferi* nucleic acids via PCR).

Where diagnosis of a disease state related to infection with *Borrelia* has already been made, the present invention is useful for monitoring progression or regression of the disease state whereby patients exhibiting enhanced *Borrelia* gene expression will experience a worse clinical outcome relative to patients expressing these gene(s) at a lower level.

By "biological sample" is intended any biological sample obtained from an animal, cell line, tissue culture, or other source which contains *Borrelia* polypeptide, mRNA, or DNA. Biological samples include body fluids (such as saliva, blood, plasma, urine, mucus, synovial fluid, etc.) tissues (such as muscle, skin, and cartilage) and any other biological source suspected of containing *Borrelia* polypeptides or nucleic acids. Methods for obtaining biological samples such as tissue are well known in the art.

The present invention is useful for detecting diseases related to *Borrelia* infections in animals. Preferred animals include monkeys, apes, cats, dogs, birds, cows, pigs, mice, horses, rabbits and humans. Particularly preferred are humans.

Total RNA can be isolated from a biological sample using any suitable technique such as the single-step guanidinium-thiocyanate-phenol-chloroform method described in Chomczynski et al. (1987) Anal. Biochem. 162:156-159. mRNA encoding *Borrelia* polypeptides having sufficient homology to the nucleic acid sequences identified in SEQ ID NOS:1-155 to allow for hybridization between complementary sequences are then assayed using any appropriate method. These include Northern blot analysis, S1 nuclease mapping, the polymerase chain reaction (PCR), reverse transcription in combination with the polymerase chain reaction (RT-PCR), and reverse transcription in combination with the ligase chain reaction (RT-LCR).

Northern blot analysis can be performed as described in Harada et al. (1990) Cell 63:303-312. Briefly, total RNA is prepared from a biological sample as described above. For the Northern blot, the RNA is denatured in an appropriate buffer (such as glyoxal/dimethyl sulfoxide/sodium phosphate buffer), subjected to agarose gel electrophoresis, and transferred onto a nitrocellulose filter. After the RNAs have been linked to the filter by a UV linker, the filter is prehybridized in a solution containing formamide, SSC, Denhardt's solution, denatured salmon sperm, SDS, and sodium phosphate buffer. A *B. burgdorferi* polynucleotide sequence shown in SEQ ID NOS:1-155 labeled according to any appropriate method (such as the ³²P-multiprimered DNA labeling system (Amersham)) is used as probe. After hybridization overnight, the filter is washed and exposed to x-ray film. DNA for use as probe according to the present invention is described in the sections above and will preferably at least 15 nucleotides in length.

S1 mapping can be performed as described in Fujita et al. (1987) Cell 49:357-367. To prepare probe DNA for use in S1 mapping, the sense strand of an above-described *B. burgdorferi* DNA sequence of the present invention is used as a template to synthesize labeled antisense DNA. The antisense DNA can then be digested using an appropriate restriction endonuclease to generate further DNA probes of a desired length. Such antisense probes are useful for visualizing protected bands corresponding to the target mRNA (i.e., mRNA encoding *Borrelia* polypeptides).

Levels of mRNA encoding *Borrelia* polypeptides are assayed, for e.g., using the RT-PCR method described in Makino et al. (1990) Technique 2:295-301. By this method, the radioactivities of the "amplicons" in the polyacrylamide gel bands are linearly related to the initial concentration of the target mRNA. Briefly, this method involves adding total RNA isolated from a biological sample in a reaction mixture containing a RT primer and appropriate buffer. After incubating for primer annealing, the mixture can be supplemented with a RT buffer, dNTPs, DTT, RNase inhibitor and reverse transcriptase. After incubation to achieve reverse transcription of the RNA, the RT products are then subject to PCR using labeled primers. Alternatively, rather than labeling the primers, a labeled dNTP can be included in the PCR reaction mixture. PCR amplification can be performed in a DNA thermal cycler according to conventional techniques. After a suitable number of rounds to achieve amplification, the PCR reaction mixture is electrophoresed on a polyacrylamide gel. After drying the gel, the radioactivity of the appropriate

bands (corresponding to the mRNA encoding the *Borrelia* polypeptides of the present invention) are quantified using an imaging analyzer. RT and PCR reaction ingredients and conditions, reagent and gel concentrations, and labeling methods are well known in the art. Variations on the RT-PCR method will be apparent to the skilled artisan. Other PCR methods that can detect the nucleic acid of the present invention can be found in PCR PRIMER: A LABORATORY MANUAL (C.W. Dieffenbach et al. eds., Cold Spring Harbor Lab Press, 1995).

The polynucleotides of the present invention, including both DNA and RNA, may be used to detect polynucleotides of the present invention or *Borrelia* species including *B. burgdorferi* using bio chip technology. The present invention includes both high density chip arrays (>1000 oligonucleotides per cm²) and low density chip arrays (<1000 oligonucleotides per cm²). Bio chips comprising arrays of polynucleotides of the present invention may be used to detect *Borrelia* species, including *B. burgdorferi*, in biological and environmental samples and to diagnose an animal, including humans, with an *B. burgdorferi* or other *Borrelia* infection. The bio chips of the present invention may comprise polynucleotide sequences of other pathogens including bacteria, viral, parasitic, and fungal polynucleotide sequences, in addition to the polynucleotide sequences of the present invention, for use in rapid differential pathogenic detection and diagnosis. The bio chips can also be used to monitor an *B. burgdorferi* or other *Borrelia* infections and to monitor the genetic changes (deletions, insertions, mismatches, etc.) in response to drug therapy in the clinic and drug development in the laboratory. The bio chip technology comprising arrays of polynucleotides of the present invention may also be used to simultaneously monitor the expression of a multiplicity of genes, including those of the present invention. The polynucleotides used to comprise a selected array may be specified in the same manner as for the fragments, i.e., by their 5' and 3' positions or length in contiguous base pairs and include from. Methods and particular uses of the polynucleotides of the present invention to detect *Borrelia* species, including *B. burgdorferi*, using bio chip technology include those known in the art and those of: U.S. Patent Nos. 5510270, 5545531, 5445934, 5677195, 5532128, 5556752, 5527681, 5451683, 5424186, 5607646, 5658732 and World Patent Nos. WO/9710365, WO/9511995, WO/9743447, WO/9535505, each incorporated herein in their entireties.

Biosensors using the polynucleotides of the present invention may also be used to detect, diagnose, and monitor *B. burgdorferi* or other *Borrelia* species and infections thereof. Biosensors using the polynucleotides of the present invention may also be used to detect particular polynucleotides of the present invention. Biosensors using the polynucleotides of the present invention may also be used to monitor the genetic changes (deletions, insertions, mismatches, etc.) in response to drug therapy in the clinic and drug development in the laboratory. Methods and particular uses of the polynucleotides of the present invention to detect *Borrelia* species, including *B. burgdorferi*, using biosensors include those known in the art and those of: U.S. Patent Nos. 5721102, 5658732, 5631170, and World Patent Nos. WO97/35011, WO/9720203, each incorporated herein in their entireties.

Thus, the present invention includes both bio chips and biosensors comprising polynucleotides of the present invention and methods of their use.

Assaying *Borrelia* polypeptide levels in a biological sample can occur using any art-known method, such as antibody-based techniques. For example, *Borrelia* polypeptide expression in tissues can be studied with classical immunohistological methods. In these, the specific recognition is provided by the primary antibody (polyclonal or monoclonal) but the secondary detection system can utilize fluorescent, enzyme, or other conjugated secondary antibodies. As a result, an immunohistological staining of tissue section for pathological examination is obtained. Tissues can also be extracted, e.g., with urea and neutral detergent, for the liberation of *Borrelia* polypeptides for Western-blot or dot/slot assay. See, e.g., Jalkanen, M. et al. (1985) J. Cell. Biol. 101:976-985; Jalkanen, M. et al. (1987) J. Cell. Biol. 105:3087-3096. In this technique, which is based on the use of cationic solid phases, quantitation of a *Borrelia* polypeptide can be accomplished using an isolated *Borrelia* polypeptide as a standard. This technique can also be applied to body fluids.

Other antibody-based methods useful for detecting *Borrelia* polypeptide gene expression include immunoassays, such as the ELISA and the radioimmunoassay (RIA). For example, a *Borrelia* polypeptide-specific monoclonal antibodies can be used both as an immunoabsorbent and as an enzyme-labeled probe to detect and quantify a *Borrelia* polypeptide. The amount of a *Borrelia* polypeptide present in the sample can be calculated by reference to the amount present in a standard preparation using a linear regression computer algorithm. Such an ELISA is described in Iacobelli et al. (1988) Breast Cancer Research and Treatment 11:19-30. In another ELISA assay, two distinct specific monoclonal antibodies can be used to detect *Borrelia* polypeptides in a body fluid. In this assay, one of the antibodies is used as the immunoabsorbent and the other as the enzyme-labeled probe.

The above techniques may be conducted essentially as a "one-step" or "two-step" assay. The "one-step" assay involves contacting the *Borrelia* polypeptide with immobilized antibody and, without washing, contacting the mixture with the labeled antibody. The "two-step" assay involves washing before contacting the mixture with the labeled antibody. Other conventional methods may also be employed as suitable. It is usually desirable to immobilize one component of the assay system on a support, thereby allowing other components of the system to be brought into contact with the component and readily removed from the sample. Variations of the above and other immunological methods included in the present invention can also be found in Harlow et al., ANTIBODIES: A LABORATORY MANUAL, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988).

Suitable enzyme labels include, for example, those from the oxidase group, which catalyze the production of hydrogen peroxide by reacting with substrate. Glucose oxidase is particularly preferred as it has good stability and its substrate (glucose) is readily available. Activity of an oxidase label may be assayed by measuring the concentration of hydrogen peroxide formed by the enzyme-labeled antibody/substrate reaction. Besides enzymes, other suitable

labels include radioisotopes, such as iodine (^{125}I , ^{121}I), carbon (^{14}C), sulphur (^{35}S), tritium (^3H), indium (^{112}In), and technetium ($^{99\text{m}}\text{Tc}$), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

Further suitable labels for the *Borrelia* polypeptide-specific antibodies of the present invention are provided below. Examples of suitable enzyme labels include malate dehydrogenase, *Borrelia* nuclease, delta-5-steroid isomerase, yeast-alcohol dehydrogenase, alpha-glycerol phosphate dehydrogenase, triose phosphate isomerase, peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase, and acetylcholine esterase.

Examples of suitable radioisotopic labels include ^3H , ^{111}In , ^{125}I , ^{131}I , ^{32}P , ^{35}S , ^{14}C , ^{51}Cr , ^{57}To , ^{58}Co , ^{59}Fe , ^{75}Se , ^{152}Eu , ^{90}Y , ^{67}Cu , ^{217}Ci , ^{211}At , ^{212}Pb , ^{47}Sc , ^{109}Pd , etc. ^{111}In is a preferred isotope where *in vivo* imaging is used since it avoids the problem of dehalogenation of the ^{125}I or ^{131}I -labeled monoclonal antibody by the liver. In addition, this radionucleotide has a more favorable gamma emission energy for imaging. See, e.g., Perkins et al. (1985) Eur. J. Nucl. Med. 10:296-301; Carasquillo et al. (1987) J. Nucl. Med. 28:281-287. For example, ^{111}In coupled to monoclonal antibodies with 1-(P-isothiocyanatobenzyl)-DPTA has shown little uptake in non-tumors tissues, particularly the liver, and therefore enhances specificity of tumor localization. See, Esteban et al. (1987) J. Nucl. Med. 28:861-870.

Examples of suitable non-radioactive isotopic labels include ^{157}Gd , ^{55}Mn , ^{162}Dy , ^{52}Tr , and ^{56}Fe .

Examples of suitable fluorescent labels include an ^{152}Eu label, a fluorescein label, an isothiocyanate label, a rhodamine label, a phycoerythrin label, a phycocyanin label, an allophycocyanin label, an o-phthaldehyde label, and a fluorescamine label.

Examples of suitable toxin labels include, *Pseudomonas* toxin, diphtheria toxin, ricin, and cholera toxin.

Examples of chemiluminescent labels include a luminal label, an isoluminal label, an aromatic acridinium ester label, an imidazole label, an acridinium salt label, an oxalate ester label, a luciferin label, a luciferase label, and an aequorin label.

Examples of nuclear magnetic resonance contrasting agents include heavy metal nuclei such as Gd, Mn, and iron.

Typical techniques for binding the above-described labels to antibodies are provided by Kennedy et al. (1976) Clin. Chim. Acta 70:1-31, and Schurs et al. (1977) Clin. Chim. Acta 81:1-40. Coupling techniques mentioned in the latter are the glutaraldehyde method, the periodate method, the dimaleimide method, the m-maleimidobenzyl-N-hydroxy-succinimide ester method, all of which methods are incorporated by reference herein.

In a related aspect, the invention includes a diagnostic kit for use in screening serum containing antibodies specific against *B. burgdorferi* infection. Such a kit may include an isolated *B. burgdorferi* antigen comprising an epitope which is specifically immunoreactive with at least one anti-*B. burgdorferi* antibody. Such a kit also includes means for detecting the

binding of said antibody to the antigen. In specific embodiments, the kit may include a recombinantly produced or chemically synthesized peptide or polypeptide antigen. The peptide or polypeptide antigen may be attached to a solid support.

In a more specific embodiment, the detecting means of the above-described kit includes a solid support to which said peptide or polypeptide antigen is attached. Such a kit may also include a non-attached reporter-labeled anti-human antibody. In this embodiment, binding of the antibody to the *B. burgdorferi* antigen can be detected by binding of the reporter labeled antibody to the anti-*B. burgdorferi* polypeptide antibody.

In a related aspect, the invention includes a method of detecting *B. burgdorferi* infection in a subject. This detection method includes reacting a body fluid, preferably serum, from the subject with an isolated *B. burgdorferi* antigen, and examining the antigen for the presence of bound antibody. In a specific embodiment, the method includes a polypeptide antigen attached to a solid support, and serum is reacted with the support. Subsequently, the support is reacted with a reporter-labeled anti-human antibody. The support is then examined for the presence of reporter-labeled antibody.

The solid surface reagent employed in the above assays and kits is prepared by known techniques for attaching protein material to solid support material, such as polymeric beads, dip sticks, 96-well plates or filter material. These attachment methods generally include non-specific adsorption of the protein to the support or covalent attachment of the protein, typically through a free amine group, to a chemically reactive group on the solid support, such as an activated carboxyl, hydroxyl, or aldehyde group. Alternatively, streptavidin coated plates can be used in conjunction with biotinylated antigen(s).

The polypeptides and antibodies of the present invention, including fragments thereof, may be used to detect *Borrelia* species including *B. burgdorferi* using bio chip and biosensor technology. Bio chip and biosensors of the present invention may comprise the polypeptides of the present invention to detect antibodies, which specifically recognize *Borrelia* species, including *B. burgdorferi*. Bio chip and biosensors of the present invention may also comprise antibodies which specifically recognize the polypeptides of the present invention to detect *Borrelia* species, including *B. burgdorferi* or specific polypeptides of the present invention. Bio chips or biosensors comprising polypeptides or antibodies of the present invention may be used to detect *Borrelia* species, including *B. burgdorferi*, in biological and environmental samples and to diagnose an animal, including humans, with an *B. burgdorferi* or other *Borrelia* infection. Thus, the present invention includes both bio chips and biosensors comprising polypeptides or antibodies of the present invention and methods of their use.

The bio chips of the present invention may further comprise polypeptide sequences of other pathogens including bacteria, viral, parasitic, and fungal polypeptide sequences, in addition to the polypeptide sequences of the present invention, for use in rapid differential pathogenic detection and diagnosis. The bio chips of the present invention may further comprise antibodies or fragments thereof specific for other pathogens including bacteria, viral, parasitic, and fungal

polypeptide sequences, in addition to the antibodies or fragments thereof of the present invention, for use in rapid differential pathogenic detection and diagnosis. The bio chips and biosensors of the present invention may also be used to monitor an *B. burgdorferi* or other *Borrelia* infection and to monitor the genetic changes (amino acid deletions, insertions, substitutions, etc.) in response to drug therapy in the clinic and drug development in the laboratory. The bio chip and biosensors comprising polypeptides or antibodies of the present invention may also be used to simultaneously monitor the expression of a multiplicity of polypeptides, including those of the present invention. The polypeptides used to comprise a bio chip or biosensor of the present invention may be specified in the same manner as for the fragments, i.e., by their N-terminal and C-terminal positions or length in contiguous amino acid residue. Methods and particular uses of the polypeptides and antibodies of the present invention to detect *Borrelia* species, including *B. burgdorferi*, or specific polypeptides using bio chip and biosensor technology include those known in the art, those of the U.S. Patent Nos. and World Patent Nos. listed above for bio chips and biosensors using polynucleotides of the present invention, and those of: U.S. Patent Nos. 5658732, 5135852, 5567301, 5677196, 5690894 and World Patent Nos. WO9729366, WO9612957, each incorporated herein in their entireties.

5. Screening Assay for Binding Agents

Using the isolated proteins of the present invention, the present invention further provides methods of obtaining and identifying agents which bind to a protein encoded by one of the ORFs of the present invention or to one of the fragments and the *Borrelia burgdorferi* fragment and contigs herein described.

In general, such methods comprise steps of:

- (a) contacting an agent with an isolated protein encoded by one of the ORFs of the present invention, or an isolated fragment of the *Borrelia burgdorferi* genome; and
- (b) determining whether the agent binds to said protein or said fragment.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention.

Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like capable of binding to a specific peptide sequence in order to generate rationally designed anti-peptide peptides, for example see Hurby *et al.*, "Application of Synthetic Peptides: Antisense Peptides," in *Synthetic*

Peptides, A User's Guide, W. H. Freeman, NY (1992), pp. 289-307, and Kaspczak *et al.*, *Biochemistry* 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control.

One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods usually contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee *et al.*, *Nucl. Acids Res.* 6:3073 (1979); Cooney *et al.*, *Science* 241:456 (1988); and Dervan *et al.*, *Science* 251:1360 (1991)) or to the mRNA itself (antisense - Okano, *J. Neurochem.* 56:560 (1991); *Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression*, CRC Press, Boca Raton, FL (1988)). Triple helix- formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention can be used to design antisense and triple helix-forming oligonucleotides, and other DNA binding agents.

6. Pharmaceutical Compositions and Vaccines

The present invention further provides pharmaceutical agents which can be used to modulate the growth or pathogenicity of *Borrelia burgdorferi*, or another related organism, *in vivo* or *in vitro*. As used herein, a "pharmaceutical agent" is defined as a composition of matter which can be formulated using known techniques to provide a pharmaceutical compositions. As used herein, the "pharmaceutical agents of the present invention" refers the pharmaceutical agents which are derived from the proteins encoded by the ORFs of the present invention or are agents which are identified using the herein described assays.

As used herein, a pharmaceutical agent is said to "modulate the growth pathogenicity of *Borrelia burgdorferi* or a related organism, *in vivo* or *in vitro*," when the agent reduces the rate of growth, rate of division, or viability of the organism in question. The pharmaceutical agents of the present invention can modulate the growth or pathogenicity of an organism in many fashions, although an understanding of the underlying mechanism of action is not needed to practice the use of the pharmaceutical agents of the present invention. Some agents will modulate the growth by binding to an important protein thus blocking the biological activity of the protein, while other agents may bind to a component of the outer surface of the organism blocking attachment or

rendering the organism more prone to act the bodies nature immune system. Alternatively, the agent may comprise a protein encoded by one of the ORFs of the present invention and serve as a vaccine. The development and use of a vaccine based on outer membrane components are well known in the art.

5 As used herein, a "related organism" is a broad term which refers to any organism whose growth can be modulated by one of the pharmaceutical agents of the present invention. In general, such an organism will contain a homolog of the protein which is the target of the pharmaceutical agent or the protein used as a vaccine. As such, related organisms do not need to be bacterial but may be fungal or viral pathogens.

10 The pharmaceutical agents and compositions of the present invention may be administered in a convenient manner, such as by the oral, topical, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes. The pharmaceutical compositions are administered in an amount which is effective for treating and/or prophylaxis of the specific indication. In general, they are administered in an amount of at least about 1 mg/kg body weight
15 and in most cases they will be administered in an amount not in excess of about 1 g/kg body weight per day. In most cases, the dosage is from about 0.1 mg/kg to about 10 g/kg body weight daily, taking into account the routes of administration, symptoms, *etc.*

The agents of the present invention can be used in native form or can be modified to form a chemical derivative. As used herein, a molecule is said to be a "chemical derivative" of another
20 molecule when it contains additional chemical moieties not normally a part of the molecule. Such moieties may improve the molecule's solubility, absorption, biological half life, *etc.* The moieties may alternatively decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, *etc.* Moieties capable of mediating such effects are disclosed in, among other sources, REMINGTON'S PHARMACEUTICAL SCIENCES (1980)
25 cited elsewhere herein.

For example, such moieties may change an immunological character of the functional derivative, such as affinity for a given antibody. Such changes in immunomodulation activity are measured by the appropriate assay, such as a competitive type immunoassay. Modifications of such protein properties as redox or thermal stability, biological half-life, hydrophobicity,
30 susceptibility to proteolytic degradation or the tendency to aggregate with carriers or into multimers also may be effected in this way and can be assayed by methods well known to the skilled artisan.

The therapeutic effects of the agents of the present invention may be obtained by providing the agent to a patient by any suitable means (*e.g.*, inhalation, intravenously,
35 intramuscularly, subcutaneously, enterally, or parenterally). It is preferred to administer the agent of the present invention so as to achieve an effective concentration within the blood or tissue in which the growth of the organism is to be controlled. To achieve an effective blood concentration, the preferred method is to administer the agent by injection. The administration may be by continuous infusion, or by single or multiple injections.

In providing a patient with one of the agents of the present invention, the dosage of the administered agent will vary depending upon such factors as the patient's age, weight, height, sex, general medical condition, previous medical history, *etc.* In general, it is desirable to provide the recipient with a dosage of agent which is in the range of from about 1 pg/kg to 10 mg/kg (body weight of patient), although a lower or higher dosage may be administered. The therapeutically effective dose can be lowered by using combinations of the agents of the present invention or another agent.

As used herein, two or more compounds or agents are said to be administered "in combination" with each other when either (1) the physiological effects of each compound, or (2) the serum concentrations of each compound can be measured at the same time. The composition of the present invention can be administered concurrently with, prior to, or following the administration of the other agent.

The agents of the present invention are intended to be provided to recipient subjects in an amount sufficient to decrease the rate of growth (as defined above) of the target organism.

The administration of the agent(s) of the invention may be for either a "prophylactic" or "therapeutic" purpose. When provided prophylactically, the agent(s) are provided in advance of any symptoms indicative of the organisms growth. The prophylactic administration of the agent(s) serves to prevent, attenuate, or decrease the rate of onset of any subsequent infection. When provided therapeutically, the agent(s) are provided at (or shortly after) the onset of an indication of infection. The therapeutic administration of the compound(s) serves to attenuate the pathological symptoms of the infection and to increase the rate of recovery.

The agents of the present invention are administered to a subject, such as a mammal, or a patient, in a pharmaceutically acceptable form and in a therapeutically effective concentration. A composition is said to be "pharmacologically acceptable" if its administration can be tolerated by a recipient patient. Such an agent is said to be administered in a "therapeutically effective amount" if the amount administered is physiologically significant. An agent is physiologically significant if its presence results in a detectable change in the physiology of a recipient patient.

The agents of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby these materials, or their functional derivatives, are combined in a mixture with a pharmaceutically acceptable carrier vehicle. Suitable vehicles and their formulation, inclusive of other human proteins, *e.g.*, human serum albumin, are described, for example, in REMINGTON'S PHARMACEUTICAL SCIENCES, 16th Ed., Osol, A., Ed., Mack Publishing, Easton PA (1980). In order to form a pharmaceutically acceptable composition suitable for effective administration, such compositions will contain an effective amount of one or more of the agents of the present invention, together with a suitable amount of carrier vehicle.

Additional pharmaceutical methods may be employed to control the duration of action. Control release preparations may be achieved through the use of polymers to complex or absorb one or more of the agents of the present invention. The controlled delivery may be effectuated by

a variety of well known techniques, including formulation with macromolecules such as, for example, polyesters, polyamino acids, polyvinyl, pyrrolidone, ethylenevinylacetate, methylcellulose, carboxymethylcellulose, or protamine, sulfate, adjusting the concentration of the macromolecules and the agent in the formulation, and by appropriate use of methods of incorporation, which can be manipulated to effectuate a desired time course of release. Another possible method to control the duration of action by controlled release preparations is to incorporate agents of the present invention into particles of a polymeric material such as polyesters, polyamino acids, hydrogels, poly(lactic acid) or ethylene vinylacetate copolymers. Alternatively, instead of incorporating these agents into polymeric particles, it is possible to entrap these materials in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization with, for example, hydroxymethylcellulose or gelatine-microcapsules and poly(methylmethacrylate) microcapsules, respectively, or in colloidal drug delivery systems, for example, liposomes, albumin microspheres, microemulsions, nanoparticles, and nanocapsules or in macroemulsions. Such techniques are disclosed in REMINGTON'S PHARMACEUTICAL SCIENCES (1980).

The invention further provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration.

In addition, the agents of the present invention may be employed in conjunction with other therapeutic compounds.

7. Shot-Gun Approach to Megabase DNA Sequencing

The present invention further demonstrates that a large sequence can be sequenced using a random shotgun approach. This procedure, described in detail in the examples that follow, has eliminated the up front cost of isolating and ordering overlapping or contiguous subclones prior to the start of the sequencing protocols.

Certain aspects of the present invention are described in greater detail in the examples that follow. The examples are provided by way of illustration. Other aspects and embodiments of the present invention are contemplated by the inventors, as will be clear to those of skill in the art from reading the present disclosure.

ILLUSTRATIVE EXAMPLES

LIBRARIES AND SEQUENCING

1. Shotgun Sequencing Probability Analysis

The overall strategy for a shotgun approach to whole genome sequencing follows from the Lander and Waterman (Landerman and Waterman, *Genomics* 2:231 (1988)) application of the equation for the Poisson distribution. According to this treatment, the probability, P_0 , that any given base in a sequence of size L , in nucleotides, is not sequenced after a certain amount, n , in nucleotides, of random sequence has been determined can be calculated by the equation $P_0 = e^{-m}$, where m is L/n , the fold coverage. For instance, for a genome of 2.8 Mb, $m=1$ when 2.8 Mb of sequence has been randomly generated (1X coverage). At that point, $P_0 = e^{-1} = 0.37$. The probability that any given base has not been sequenced is the same as the probability that any region of the whole sequence L has not been determined and, therefore, is equivalent to the fraction of the whole sequence that has yet to be determined. Thus, at one-fold coverage, approximately 37% of a polynucleotide of size L , in nucleotides has not been sequenced. When 14 Mb of sequence has been generated, coverage is 5X for a 2.8 Mb and the unsequenced fraction drops to .0067 or 0.67%. 5X coverage of a 2.8 Mb sequence can be attained by sequencing approximately 17,000 random clones from both insert ends with an average sequence read length of 410 bp.

Similarly, the total gap length, G , is determined by the equation $G = Le^{-m}$, and the average gap size, g , follows the equation, $g = L/n$. Thus, 5X coverage leaves about 240 gaps averaging about 82 bp in size in a sequence of a polynucleotide 2.8 Mb long.

The treatment above is essentially that of Lander and Waterman, *Genomics* 2: 231 (1988).

2. Random Library Construction

In order to approximate the random model described above during actual sequencing, a nearly ideal library of cloned genomic fragments is required. The following library construction procedure was developed to achieve this end.

Borrelia burgdorferi DNA is prepared by phenol extraction. A mixture containing 200 µg DNA in 1.0 ml of 300 mM sodium acetate, 10 mM Tris-HCl, 1 mM Na-EDTA, 50% glycerol is processed through a nebulizer (IPI Medical Products) with a stream of nitrogen adjusted to 35 Kpa for 2 minutes. The sonicated DNA is ethanol precipitated and redissolved in 500 µl TE buffer.

To create blunt-ends, a 100 µl aliquot of the resuspended DNA is digested with 5 units of BAL31 nuclease (New England BioLabs) for 10 min at 30°C in 200 µl BAL31 buffer. The digested DNA is phenol-extracted, ethanol-precipitated, redissolved in 100 µl TE buffer, and then size-fractionated by electrophoresis through a 1.0% low melting temperature agarose gel. The section containing DNA fragments 1.6-2.0 kb in size is excised from the gel, and the LGT agarose is melted and the resulting solution is extracted with phenol to separate the agarose from the DNA. DNA is ethanol precipitated and redissolved in 20 µl of TE buffer for ligation to vector.

A two-step ligation procedure is used to produce a plasmid library with 97% inserts, of which >99% were single inserts. The first ligation mixture (50 μ l) contains 2 μ g of DNA fragments, 2 μ g pUC18 DNA (Pharmacia) cut with SmaI and dephosphorylated with bacterial alkaline phosphatase, and 10 units of T4 ligase (GIBCO/BRL) and is incubated at 14°C for 4 hr.

- 5 The ligation mixture then is phenol extracted and ethanol precipitated, and the precipitated DNA is dissolved in 20 μ l TE buffer and electrophoresed on a 1.0% low melting agarose gel. Discrete bands in a ladder are visualized by ethidium bromide-staining and UV illumination and identified by size as insert (I), vector (v), v+I, v+2i, v+3i, etc. The portion of the gel containing v+I DNA is excised and the v+I DNA is recovered and resuspended into 20 μ l TE. The v+I DNA then is
- 10 blunt-ended by T4 polymerase treatment for 5 min. at 37°C in a reaction mixture (50 μ l) containing the v+I linears, 500 μ M each of the 4 dNTPs, and 9 units of T4 polymerase (New England BioLabs), under recommended buffer conditions. After phenol extraction and ethanol precipitation the repaired v+I linears are dissolved in 20 μ l TE. The final ligation to produce circles is carried out in a 50 μ l reaction containing 5 μ l of v+I linears and 5 units of T4 ligase at
- 15 14°C overnight. After 10 min. at 70°C the following day, the reaction mixture is stored at -20°C.

This two-stage procedure results in a molecularly random collection of single-insert plasmid recombinants with minimal contamination from double-insert chimeras (<1%) or free vector (<3%).

- Since deviation from randomness can arise from propagation the DNA in the host, *E. coli*
- 20 host cells deficient in all recombination and restriction functions (A. Greener, *Strategies* 3 (1):5 (1990)) are used to prevent rearrangements, deletions, and loss of clones by restriction. Furthermore, transformed cells are plated directly on antibiotic diffusion plates to avoid the usual broth recovery phase which allows multiplication and selection of the most rapidly growing cells.

- Plating is carried out as follows. A 100 μ l aliquot of Epicurian Coli SURE II
- 25 Supercompetent Cells (Stratagene 200152) is thawed on ice and transferred to a chilled Falcon 2059 tube on ice. A 1.7 μ l aliquot of 1.42 M beta-mercaptoethanol is added to the aliquot of cells to a final concentration of 25 mM. Cells are incubated on ice for 10 min. A 1 μ l aliquot of the final ligation is added to the cells and incubated on ice for 30 min. The cells are heat pulsed for 30 sec. at 42°C and placed back on ice for 2 min. The outgrowth period in liquid culture is
- 30 eliminated from this protocol in order to minimize the preferential growth of any given transformed cell. Instead the transformation mixture is plated directly on a nutrient rich SOB plate containing a 5 ml bottom layer of SOB agar (5% SOB agar: 20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 1.5% Difco Agar per liter of media). The 5 ml bottom layer is supplemented with 0.4 ml of 50 mg/ml ampicillin per 100 ml SOB agar. The 15 ml top layer of SOB agar is
- 35 supplemented with 1 ml X-Gal (2%), 1 ml MgCl₂ (1 M), and 1 ml MgSO₄/100 ml SOB agar. The 15 ml top layer is poured just prior to plating. Our titer is approximately 100 colonies/10 μ l aliquot of transformation.

All colonies are picked for template preparation regardless of size. Thus, only clones lost due to "poison" DNA or deleterious gene products are deleted from the library, resulting in a slight increase in gap number over that expected.

3. Random DNA Sequencing

High quality double stranded DNA plasmid templates are prepared using a "boiling bead" method developed in collaboration with Advanced Genetic Technology Corp. (Gaithersburg, MD) (Adams *et al.*, *Science* 252:1651 (1991); Adams *et al.*, *Nature* 355:632 (1992)). Plasmid preparation is performed in a 96-well format for all stages of DNA preparation from bacterial growth through final DNA purification. Template concentration is determined using Hoechst Dye and a Millipore Cytofluor. DNA concentrations are not adjusted, but low-yielding templates are identified where possible and not sequenced.

Templates are also prepared from two *Borrelia burgdorferi* lambda genomic libraries. An amplified library is constructed in the vector Lambda GEM-12 (Promega) and an unamplified library is constructed in Lambda DASH II (Stratagene). In particular, for the unamplified lambda library, *Borrelia burgdorferi* DNA (> 100 kb) is partially digested in a reaction mixture (200 ul) containing 50 µg DNA, 1X Sau3AI buffer, 20 units Sau3AI for 6 min. at 23°C. The digested DNA was phenol-extracted and electrophoresed on a 0.5% low melting agarose gel at 2V/cm for 7 hours. Fragments from 15 to 25 kb are excised and recovered in a final volume of 6 ul. One µl of fragments is used with 1 µl of DASHII vector (Stratagene) in the recommended ligation reaction. One µl of the ligation mixture is used per packaging reaction following the recommended protocol with the Gigapack II XL Packaging Extract (Stratagene, #227711). Phage are plated directly without amplification from the packaging mixture (after dilution with 500 µl of recommended SM buffer and chloroform treatment). Yield is about 2.5×10^3 pfu/ul. The amplified library is prepared essentially as above except the lambda GEM-12 vector is used. After packaging, about 3.5×10^4 pfu are plated on the restrictive NM539 host. The lysate is harvested in 2 ml of SM buffer and stored frozen in 7% dimethylsulfoxide. The phage titer is approximately 1×10^9 pfu/ml.

Liquid lysates (100 µl) are prepared from randomly selected plaques (from the unamplified library) and template is prepared by long-range PCR using T7 and T3 vector-specific primers.

Sequencing reactions are carried out on plasmid and/or PCR templates using the AB Catalyst LabStation with Applied Biosystems PRISM Ready Reaction Dye Primer Cycle Sequencing Kits for the M13 forward (M13-21) and the M13 reverse (M13RP1) primers (Adams *et al.*, *Nature* 368:474 (1994)). Dye terminator sequencing reactions are carried out on the lambda templates on a Perkin-Elmer 9600 Thermocycler using the Applied Biosystems Ready Reaction Dye Terminator Cycle Sequencing kits. T7 and SP6 primers are used to sequence the ends of the inserts from the Lambda GEM-12 library and T7 and T3 primers are used to sequence the ends of the inserts from the Lambda DASH II library. Sequencing reactions are performed

by eight individuals using an average of fourteen AB 373 DNA Sequencers per day. All sequencing reactions are analyzed using the Stretch modification of the AB 373, primarily using a 34 cm well-to-read distance. The overall sequencing success rate very approximately is about 85% for M13-21 and M13RP1 sequences and 65% for dye-terminator reactions. The average usable read length is 485 bp for M13-21 sequences, 445bp for M13RP1 sequences, and 375 bp for dye-terminator reactions.

Richards *et al.*, Chapter 28 in AUTOMATED DNA SEQUENCING AND ANALYSIS, M. D. Adams, C. Fields, J. C. Venter, Eds., Academic Press, London, (1994) described the value of using sequence from both ends of sequencing templates to facilitate ordering of contigs in shotgun assembly projects of lambda and cosmid clones. We balance the desirability of both-end sequencing (including the reduced cost of lower total number of templates) against shorter read-lengths for sequencing reactions performed with the M13RP1 (reverse) primer compared to the M13-21 (forward) primer. Approximately one-half of the templates are sequenced from both ends. Random reverse sequencing reactions are done based on successful forward sequencing reactions. Some M13RP1 sequences are obtained in a semi-directed fashion: M13-21: sequences pointing outward at the ends of contigs are chosen for M13RP1 sequencing in an effort to specifically order contigs.

4. Protocol for Automated Cycle Sequencing

The sequencing is carried out using ABI Catalyst robots and AB 373 Automated DNA Sequencers. The Catalyst robot is a publicly available sophisticated pipetting and temperature control robot which has been developed specifically for DNA sequencing reactions. The Catalyst combines pre-aliquoted templates and reaction mixes consisting of deoxy- and dideoxynucleotides, the thermostable Taq DNA polymerase, fluorescently-labelled sequencing primers, and reaction buffer. Reaction mixes and templates are combined in the wells of an aluminum 96-well thermocycling plate. Thirty consecutive cycles of linear amplification (*i.e.*, one primer synthesis) steps are performed including denaturation, annealing of primer and template, and extension; *i.e.*, DNA synthesis. A heated lid with rubber gaskets on the thermocycling plate prevents evaporation without the need for an oil overlay.

Two sequencing protocols are used: one for dye-labelled primers and a second for dye-labelled dideoxy chain terminators. The shotgun sequencing involves use of four dye-labelled sequencing primers, one for each of the four terminator nucleotide. Each dye-primer is labelled with a different fluorescent dye, permitting the four individual reactions to be combined into one lane of the 373 DNA Sequencer for electrophoresis, detection, and base-calling. ABI currently supplies pre-mixed reaction mixes in bulk packages containing all the necessary non-template reagents for sequencing. Sequencing can be done with both plasmid and PCR- generated templates with both dye-primers and dye- terminators with approximately equal fidelity, although plasmid templates generally give longer usable sequences.

Thirty-two reactions are loaded per AB373 Sequencer each day, for a total of 960 samples. Electrophoresis is run overnight following the manufacturer's protocols, and the data is collected for twelve hours. Following electrophoresis and fluorescence detection, the ABI 373 performs automatic lane tracking and base-calling. The lane-tracking is confirmed visually. Each sequence electropherogram (or fluorescence lane trace) is inspected visually and assessed for quality. Trailing sequences of low quality are removed and the sequence itself is loaded via software to a Sybase database (archived daily to 8mm tape). Leading vector polylinker sequence is removed automatically by a software program. Average edited lengths of sequences from the standard ABI 373 are around 400 bp and depend mostly on the quality of the template used for the sequencing reaction. ABI 373 Sequencers converted to Stretch Liners provide a longer electrophoresis path prior to fluorescence detection and increase the average number of usable bases to 500-600 bp.

INFORMATICS

1. Data Management

A number of information management systems for a large-scale sequencing lab have been developed. (For review see, for instance, Kerlavage *et al.*, *Proceedings of the Twenty-Sixth Annual Hawaii International Conference on System Sciences*, IEEE Computer Society Press, Washington D. C., 585 (1993)) The system used to collect and assemble the sequence data was developed using the Sybase relational database management system and was designed to automate data flow wherever possible and to reduce user error. The database stores and correlates all information collected during the entire operation from template preparation to final analysis of the genome. Because the raw output of the ABI 373 Sequencers was based on a Macintosh platform and the data management system chosen was based on a Unix platform, it was necessary to design and implement a variety of multi-user, client-server applications which allow the raw data as well as analysis results to flow seamlessly into the database with a minimum of user effort.

2. Assembly

An assembly engine (TIGR Assembler) developed for the rapid and accurate assembly of thousands of sequence fragments was employed to generate contigs. The TIGR assembler simultaneously clusters and assembles fragments of the genome. In order to obtain the speed necessary to assemble more than 104 fragments, the algorithm builds a hash table of 12 bp oligonucleotide subsequences to generate a list of potential sequence fragment overlaps. The number of potential overlaps for each fragment determines which fragments are likely to fall into repetitive elements. Beginning with a single seed sequence fragment, TIGR Assembler extends the current contig by attempting to add the best matching fragment based on oligonucleotide content. The contig and candidate fragment are aligned using a modified version of the Smith-Waterman algorithm which provides for optimal gapped alignments (Waterman, M. S., *Methods*

in *Enzymology* 164:765 (1988)). The contig is extended by the fragment only if strict criteria for the quality of the match are met. The match criteria include the minimum length of overlap, the maximum length of an unmatched end, and the minimum percentage match. These criteria are automatically lowered by the algorithm in regions of minimal coverage and raised in regions with a possible repetitive element. The number of potential overlaps for each fragment determines which fragments are likely to fall into repetitive elements. Fragments representing the boundaries of repetitive elements and potentially chimeric fragments are often rejected based on partial mismatches at the ends of alignments and excluded from the current contig. TIGR Assembler is designed to take advantage of clone size information coupled with sequencing from both ends of each template. It enforces the constraint that sequence fragments from two ends of the same template point toward one another in the contig and are located within a certain range of base pairs (definable for each clone based on the known clone size range for a given library). The process resulted in 155 contigs as represented by SEQ ID NOs:1-155.

3. Identifying Genes

The predicted coding regions of the *Borrelia burgdorferi* genome were initially defined with the program GeneMark, which finds ORFs using a probabilistic classification technique. The predicted coding region sequences were used in searches against a database of all nucleotide sequences from GenBank (July, 1997), using the BLASTN search method to identify overlaps of 50 or more nucleotides with at least a 95% identity (using default parameters). Those ORFs with nucleotide sequence matches are shown in Table 1. The ORFs without such matches were translated to protein sequences and compared to a non-redundant database of known proteins generated by combining the Swiss-prot, PIR and GenPept databases. ORFs that matched a database protein with BLASTP probability less than or equal to 0.01 are shown in Table 2. The table also lists assigned functions based on the closest match in the databases. ORFs that did not match protein or nucleotide sequences in the databases at these levels are shown in Table 3.

ILLUSTRATIVE APPLICATIONS

1. Production of an Antibody to a *Borrelia burgdorferi* Protein

Substantially pure protein or polypeptide is isolated from the transfected or transformed cells using any one of the methods known in the art. The protein can also be produced in a recombinant prokaryotic expression system, such as *E. coli*, or can be chemically synthesized.

Concentration of protein in the final preparation is adjusted, for example, by concentration on an Amicon filter device, to the level of a few micrograms/ml. Monoclonal or polyclonal antibody to the protein can then be prepared as follows.

2. Monoclonal Antibody Production by Hybridoma Fusion

Monoclonal antibody to epitopes of any of the peptides identified and isolated as described can be prepared from murine hybridomas according to the classical method of Kohler, G. and Milstein, C., *Nature* 256:495 (1975) or modifications of the methods thereof. Briefly, a mouse is repetitively inoculated with a few micrograms of the selected protein over a period of a few weeks. The mouse is then sacrificed, and the antibody producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and the excess unfused cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). The successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate where growth of the culture is continued. Antibody-producing clones are identified by detection of antibody in the supernatant fluid of the wells by immunoassay procedures, such as ELISA, as originally described by Engvall, E., *Meth. Enzymol.* 70:419 (1980), and modified methods thereof. Selected positive clones can be expanded and their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Davis, L. *et al.*, *Basic Methods in Molecular Biology*, Elsevier, New York. Section 21-2 (1989).

3. Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogenous epitopes of a single protein can be prepared by immunizing suitable animals with the expressed protein described above, which can be unmodified or modified to enhance immunogenicity. Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. For example, small molecules tend to be less immunogenic than others and may require the use of carriers and adjuvant. Also, host animals vary in response to site of inoculations and dose, with both inadequate or excessive doses of antigen resulting in low titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appears to be most reliable. An effective immunization protocol for rabbits can be found in Vaitukaitis, J. *et al.*, *J. Clin. Endocrinol. Metab.* 33:988-991 (1971).

Booster injections can be given at regular intervals, and antiserum harvested when antibody titer thereof, as determined semi-quantitatively, for example, by double immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony, O. *et al.*, Chap. 19 in: *Handbook of Experimental Immunology*, Wier, D., ed, Blackwell (1973). Plateau concentration of antibody is usually in the range of 0.1 to 0.2 mg/ml of serum (about 12M). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher, D., Chap. 42 in: *Manual of Clinical Immunology*, second edition, Rose and Friedman, eds., Amer. Soc. For Microbiology, Washington, D. C. (1980)

Antibody preparations prepared according to either protocol are useful in quantitative immunoassays which determine concentrations of antigen-bearing substances in biological

samples; they are also used semi- quantitatively or qualitatively to identify the presence of antigen in a biological sample. In addition, antibodies are useful in various animal models of pneumococcal disease as a means of evaluating the protein used to make the antibody as a potential vaccine target or as a means of evaluating the antibody as a potential immunotherapeutic or immunoprophylactic reagent.

4. Preparation of PCR Primers and Amplification of DNA

Various fragments of the *Borrelia burgdorferi* genome, such as those of Tables 1-6 and SEQ ID NOS: 1-155 can be used, in accordance with the present invention, to prepare PCR primers for a variety of uses. The PCR primers are preferably at least 15 bases, and more preferably at least 18 bases in length. When selecting a primer sequence, it is preferred that the primer pairs have approximately the same G/C ratio, so that melting temperatures are approximately the same. The PCR primers and amplified DNA of this Example find use in the Examples that follow.

5. Isolation of a Selected DNA Clone From *B. burgdorferi*

Three approaches are used to isolate a *B. burgdorferi* clone comprising a polynucleotide of the present invention from any *B. burgdorferi* genomic DNA library. The *B. burgdorferi* strain B31PU has been deposited as a convenient source for obtaining a *B. burgdorferi* strain although a wide variety of strains *B. burgdorferi* strains can be used which are known in the art.

B. burgdorferi genomic DNA is prepared using the following method. A 20ml overnight bacterial culture grown in a rich medium (e.g., Trypticase Soy Broth, Brain Heart Infusion broth or Super broth), pelleted, washed two times with TES (30mM Tris-pH 8.0, 25mM EDTA, 50mM NaCl), and resuspended in 5ml high salt TES (2.5M NaCl). Lysostaphin is added to final concentration of approx 50ug/ml and the mixture is rotated slowly 1 hour at 37C to make protoplast cells. The solution is then placed in incubator (or place in a shaking water bath) and warmed to 55C. Five hundred micro liter of 20% sarcosyl in TES (final concentration 2%) is then added to lyse the cells. Next, guanidine HCl is added to a final concentration of 7M (3.69g in 5.5 ml). The mixture is swirled slowly at 55C for 60-90 min (solution should clear). A CsCl gradient is then set up in SW41 ultra clear tubes using 2.0ml 5.7M CsCl and overlaying with 2.85M CsCl. The gradient is carefully overlayed with the DNA-containing GuHCl solution. The gradient is spun at 30,000 rpm, 20C for 24 hr and the lower DNA band is collected. The volume is increased to 5 ml with TE buffer. The DNA is then treated with protease K (10 ug/ml) overnight at 37 C, and precipitated with ethanol. The precipitated DNA is resuspended in a desired buffer.

In the first method, a plasmid is directly isolated by screening a plasmid *B. burgdorferi* genomic DNA library using a polynucleotide probe corresponding to a polynucleotide of the present invention. Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The

oligonucleotide is labeled, for instance, with ^{32}P - γ -ATP using T4 polynucleotide kinase and purified according to routine methods. (See, e.g., Maniatis et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring, NY (1982).) The library is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art. See, e.g., Sambrook et al. *MOLECULAR CLONING: A LABORATORY MANUAL* (Cold Spring Harbor, N.Y. 2nd ed. 1989); Ausubel et al., *CURRENT PROTOCOLS IN MOLECULAR BIOLOGY* (John Wiley and Sons, N.Y. 1989). The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening. See, e.g., Sambrook et al. *MOLECULAR CLONING: A LABORATORY MANUAL* (Cold Spring Harbor, N.Y. 2nd ed. 1989); Ausubel et al., *CURRENT PROTOCOLS IN MOLECULAR BIOLOGY* (John Wiley and Sons, N.Y. 1989) or other techniques known to those of skill in the art.

Alternatively, two primers of 15-25 nucleotides derived from the 5' and 3' ends of a polynucleotide of SEQ ID NOS:1-155 are synthesized and used to amplify the desired DNA by PCR using a *B. burgdorferi* genomic DNA prep as a template. PCR is carried out under routine conditions, for instance, in 25 μl of reaction mixture with 0.5 μg of the above DNA template. A convenient reaction mixture is 1.5-5 mM MgCl_2 , 0.01% (w/v) gelatin, 20 μM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Finally, overlapping oligos of the DNA sequences of SEQ ID NOS:1-155 can be chemically synthesized and used to generate a nucleotide sequence of desired length using PCR methods known in the art.

6(a). Expression and Purification Borrelia polypeptides in E. coli

The bacterial expression vector pQE60 is used for bacterial expression of some of the polypeptide fragments of the present invention. (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311). pQE60 encodes ampicillin antibiotic resistance ("Ampr") and contains a bacterial origin of replication ("ori"), an IPTG inducible promoter, a ribosome binding site ("RBS"), six codons encoding histidine residues that allow affinity purification using nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin (QIAGEN, Inc., *supra*) and suitable single restriction enzyme cleavage sites. These elements are arranged such that an inserted DNA fragment encoding a polypeptide expresses that polypeptide with the six His residues (i.e., a "6

X His tag") covalently linked to the carboxyl terminus of that polypeptide.

The DNA sequence encoding the desired portion of a *B. burgdorferi* protein of the present invention is amplified from *B. burgdorferi* genomic DNA using PCR oligonucleotide primers which anneal to the 5' and 3' sequences coding for the portions of the *B. burgdorferi* polynucleotide shown in SEQ ID NOS:1-155. Additional nucleotides containing restriction sites to facilitate cloning in the pQE60 vector are added to the 5' and 3' sequences, respectively.

For cloning the mature protein, the 5' primer has a sequence containing an appropriate restriction site followed by nucleotides of the amino terminal coding sequence of the desired *B. burgdorferi* polynucleotide sequence in SEQ ID NOS:1-155. One of ordinary skill in the art would appreciate that the point in the protein coding sequence where the 5' and 3' primers begin may be varied to amplify a DNA segment encoding any desired portion of the complete protein shorter or longer than the mature form. The 3' primer has a sequence containing an appropriate restriction site followed by nucleotides complementary to the 3' end of the polypeptide coding sequence of SEQ ID NOS:1-155, excluding a stop codon, with the coding sequence aligned with the restriction site so as to maintain its reading frame with that of the six His codons in the pQE60 vector.

The amplified *B. burgdorferi* DNA fragment and the vector pQE60 are digested with restriction enzymes which recognize the sites in the primers and the digested DNAs are then ligated together. The *B. burgdorferi* DNA is inserted into the restricted pQE60 vector in a manner which places the *B. burgdorferi* protein coding region downstream from the IPTG-inducible promoter and in-frame with an initiating AUG and the six histidine codons.

The ligation mixture is transformed into competent *E. coli* cells using standard procedures such as those described by Sambrook et al., *supra*. *E. coli* strain M15/rep4, containing multiple copies of the plasmid pREP4, which expresses the lac repressor and confers kanamycin resistance ("Kanr"), is used in carrying out the illustrative example described herein. This strain, which is only one of many that are suitable for expressing a *B. burgdorferi* polypeptide, is available commercially (QIAGEN, Inc., *supra*). Transformants are identified by their ability to grow on LB agar plates in the presence of ampicillin and kanamycin. Plasmid DNA is isolated from resistant colonies and the identity of the cloned DNA confirmed by restriction analysis, PCR and DNA sequencing.

Clones containing the desired constructs are grown overnight ("O/N") in liquid culture in LB media supplemented with both ampicillin (100 µg/ml) and kanamycin (25 µg/ml). The O/N culture is used to inoculate a large culture, at a dilution of approximately 1:25 to 1:250. The cells are grown to an optical density at 600 nm ("OD600") of between 0.4 and 0.6. Isopropyl-β-D-thiogalactopyranoside ("IPTG") is then added to a final concentration of 1 mM to induce transcription from the lac repressor sensitive promoter, by inactivating the lacI repressor. Cells subsequently are incubated further for 3 to 4 hours. Cells then are harvested by centrifugation.

The cells are then stirred for 3-4 hours at 4°C in 6M guanidine-HCl, pH 8. The cell

debris is removed by centrifugation, and the supernatant containing the *B. burgdorferi* polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity are purified in a simple one-step procedure (for details see: The QIAexpressionist, 1995, QIAGEN, Inc., *supra*). Briefly the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the *B. burgdorferi* polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein could be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins can be eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

The polypeptide of the present invention are also prepared using a non-denaturing protein purification method. For these polypeptides, the cell pellet from each liter of culture is resuspended in 25 mls of Lysis Buffer A at 4°C (Lysis Buffer A = 50 mM Na-phosphate, 300 mM NaCl, 10 mM 2-mercaptoethanol, 10% Glycerol, pH 7.5 with 1 tablet of Complete EDTA-free protease inhibitor cocktail (Boehringer Mannheim #1873580) per 50 ml of buffer). Absorbance at 550 nm is approximately 10-20 O.D./ml. The suspension is then put through three freeze/thaw cycles from -70°C (using a ethanol-dry ice bath) up to room temperature. The cells are lysed via sonication in short 10 sec bursts over 3 minutes at approximately 80W while kept on ice. The sonicated sample is then centrifuged at 15,000 RPM for 30 minutes at 4°C. The supernatant is passed through a column containing 1.0 ml of CL-4B resin to pre-clear the sample of any proteins that may bind to agarose non-specifically, and the flow-through fraction is collected.

The pre-cleared flow-through is applied to a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (Quiagen, Inc., *supra*). Proteins with a 6 X His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure. Briefly, the supernatant is loaded onto the column in Lysis Buffer A at 4°C, the column is first washed with 10 volumes of Lysis Buffer A until the A280 of the eluate returns to the baseline. Then, the column is washed with 5 volumes of 40 mM Imidazole (92% Lysis Buffer A / 8% Buffer B) (Buffer B = 50 mM Na-Phosphate, 300 mM NaCl, 10% Glycerol, 10 mM 2-mercaptoethanol, 500 mM Imidazole, pH of the final buffer should be 7.5). The protein is eluted off of the column with a series of increasing Imidazole solutions made by adjusting the ratios of Lysis Buffer A to Buffer B. Three different concentrations are used: 3 volumes of 75 mM Imidazole, 3 volumes of

150 mM Imidazole, 5 volumes of 500 mM Imidazole. The fractions containing the purified protein are analyzed using 8 %, 10 % or 14% SDS-PAGE depending on the protein size. The purified protein is then dialyzed 2X against phosphate-buffered saline (PBS) in order to place it into an easily workable buffer. The purified protein is stored at 4°C or frozen at -80°.

5 The following alternative method may be used to purify *B. burgdorferi* expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

 Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells are harvested by continuous centrifugation at 15,000 rpm

10 (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

 The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, 15 Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 x g for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

 The resulting washed inclusion bodies are solubilized with 1.5 M guanidine 20 hydrochloride (GuHCl) for 2-4 hours. After 7000 x g centrifugation for 15 min., the pellet is discarded and the *B. burgdorferi* polypeptide-containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

 Following high speed centrifugation (30,000 x g) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of 25 buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

 To clarify the refolded *B. burgdorferi* polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area 30 (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS- 35 PAGE.

 Fractions containing the *B. burgdorferi* polypeptide are then pooled and mixed with 4

volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the *B. burgdorferi* polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant *B. burgdorferi* polypeptide exhibits greater than 95% purity after the above refolding and purification steps. No major contaminant bands are observed from Commassie blue stained 16% SDS-PAGE gel when 5 μ g of purified protein is loaded. The purified protein is also tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

6(b). Alternative Expression and Purification Borrelia polypeptides in E. coli

The vector pQE10 is alternatively used to clone and express some of the polypeptides of the present invention for use in the soft tissue and systemic infection models discussed below. The difference being such that an inserted DNA fragment encoding a polypeptide expresses that polypeptide with the six His residues (i.e., a "6 X His tag") covalently linked to the amino terminus of that polypeptide. The bacterial expression vector pQE10 (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311) was used in this example. The components of the pQE10 plasmid are arranged such that the inserted DNA sequence encoding a polypeptide of the present invention expresses the polypeptide with the six His residues (i.e., a "6 X His tag") covalently linked to the amino terminus.

The DNA sequences encoding the desired portions of a polypeptide of SEQ ID NOS:1-155 were amplified using PCR oligonucleotide primers from genomic *B. burgdorferi* DNA. The PCR primers anneal to the nucleotide sequences encoding the desired amino acid sequence of a polypeptide of the present invention. Additional nucleotides containing restriction sites to facilitate cloning in the pQE10 vector were added to the 5' and 3' primer sequences, respectively.

For cloning a polypeptide of the present invention, the 5' and 3' primers were selected to amplify their respective nucleotide coding sequences. One of ordinary skill in the art would appreciate that the point in the protein coding sequence where the 5' and 3' primers begins may be varied to amplify a DNA segment encoding any desired portion of a polypeptide of the present invention. The 5' primer was designed so the coding sequence of the 6 X His tag is aligned with the restriction site so as to maintain its reading frame with that of *B. burgdorferi* polypeptide. The 3' was designed to include an stop codon. The amplified DNA fragment was then cloned, and the protein expressed, as described above for the pQE60 plasmid.

The DNA sequences of SEQ ID NOS:1-155 encoding amino acid sequences may also be cloned and expressed as fusion proteins by a protocol similar to that described directly above, wherein the pET-32b(+) vector (Novagen, 601 Science Drive, Madison, WI 53711) is preferentially used in place of pQE10.

5 The above methods are not limited to the polypeptide fragments actually produced. The above method, like the methods below, can be used to produce either full length polypeptides or desired fragments thereof.

6(c). Alternative Expression and Purification of *Borrelia* polypeptides in

10 *E. coli*

The bacterial expression vector pQE60 is used for bacterial expression in this example (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311). However, in this example, the polypeptide coding sequence is inserted such that translation of the six His codons is prevented and, therefore, the polypeptide is produced with no 6 X His tag.

15 The DNA sequence encoding the desired portion of the *B. burgdorferi* amino acid sequence is amplified from an *B. burgdorferi* genomic DNA prep the deposited DNA clones using PCR oligonucleotide primers which anneal to the 5' and 3' nucleotide sequences corresponding to the desired portion of the *B. burgdorferi* polypeptides. Additional nucleotides containing restriction sites to facilitate cloning in the pQE60 vector are added to the 5' and 3' primer sequences.

20 For cloning a *B. burgdorferi* polypeptides of the present invention, 5' and 3' primers are selected to amplify their respective nucleotide coding sequences. One of ordinary skill in the art would appreciate that the point in the protein coding sequence where the 5' and 3' primers begin may be varied to amplify a DNA segment encoding any desired portion of a polypeptide of the present invention. The 3' and 5' primers contain appropriate restriction sites followed by nucleotides complementary to the 5' and 3' ends of the coding sequence respectively. The 3' primer is additionally designed to include an in-frame stop codon.

25 The amplified *B. burgdorferi* DNA fragments and the vector pQE60 are digested with restriction enzymes recognizing the sites in the primers and the digested DNAs are then ligated together. Insertion of the *B. burgdorferi* DNA into the restricted pQE60 vector places the *B. burgdorferi* protein coding region including its associated stop codon downstream from the IPTG-inducible promoter and in-frame with an initiating AUG. The associated stop codon prevents translation of the six histidine codons downstream of the insertion point.

30 The ligation mixture is transformed into competent *E. coli* cells using standard procedures such as those described by Sambrook et al. *E. coli* strain M15/rep4, containing multiple copies of the plasmid pREP4, which expresses the lac repressor and confers kanamycin resistance ("Kanr"), is used in carrying out the illustrative example described herein. This strain, which is only one of many that are suitable for expressing *B. burgdorferi* polypeptide, is available commercially (QIAGEN, Inc., *supra*). Transformants are identified by their ability to grow on

LB plates in the presence of ampicillin and kanamycin. Plasmid DNA is isolated from resistant colonies and the identity of the cloned DNA confirmed by restriction analysis, PCR and DNA sequencing.

Clones containing the desired constructs are grown overnight ("O/N") in liquid culture in LB media supplemented with both ampicillin (100 µg/ml) and kanamycin (25 µg/ml). The O/N culture is used to inoculate a large culture, at a dilution of approximately 1:25 to 1:250. The cells are grown to an optical density at 600 nm ("OD600") of between 0.4 and 0.6. isopropyl-b-D-thiogalactopyranoside ("IPTG") is then added to a final concentration of 1 mM to induce transcription from the *lac* repressor sensitive promoter, by inactivating the *lacI* repressor. Cells subsequently are incubated further for 3 to 4 hours. Cells then are harvested by centrifugation.

To purify the *B. burgdorferi* polypeptide, the cells are then stirred for 3-4 hours at 4°C in 6M guanidine-HCl, pH 8. The cell debris is removed by centrifugation, and the supernatant containing the *B. burgdorferi* polypeptide is dialyzed against 50 mM Na-acetate buffer pH 6, supplemented with 200 mM NaCl. Alternatively, the protein can be successfully refolded by dialyzing it against 500 mM NaCl, 20% glycerol, 25 mM Tris/HCl pH 7.4, containing protease inhibitors. After renaturation the protein can be purified by ion exchange, hydrophobic interaction and size exclusion chromatography. Alternatively, an affinity chromatography step such as an antibody column can be used to obtain pure *B. burgdorferi* polypeptide. The purified protein is stored at 4°C or frozen at -80°C.

The following alternative method may be used to purify *B. burgdorferi* polypeptides expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells are harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells were then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 x g for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 x g centrifugation for 15 min., the pellet is discarded and the *B. burgdorferi* polypeptide-containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 x g) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded *B. burgdorferi* polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the *B. burgdorferi* polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A₂₈₀ monitoring of the effluent. Fractions containing the *B. burgdorferi* polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant *B. burgdorferi* polypeptide exhibits greater than 95% purity after the above refolding and purification steps. No major contaminant bands are observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein is also tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

6(d). Cloning and Expression of *B. burgdorferi* in Other Bacteria

B. burgdorferi polypeptides can also be produced in: *B. burgdorferi* using the methods of S. Skinner et al., (1988) Mol. Microbiol. 2:289-297 or J. I. Moreno (1996) Protein Expr. Purif. 8(3):332-340; *Lactobacillus* using the methods of C. Rush et al., 1997 Appl. Microbiol. Biotechnol. 47(5):537-542; or in *Bacillus subtilis* using the methods Chang et al., U.S. Patent No. 4,952,508.

7. Cloning and Expression in COS Cells

A *B. burgdorferi* expression plasmid is made by cloning a portion of the DNA encoding a

B. burgdorferi polypeptide into the expression vector pDNAI/Amp or pDNAIII (which can be obtained from Invitrogen, Inc.). The expression vector pDNAI/amp contains: (1) an *E. coli* origin of replication effective for propagation in *E. coli* and other prokaryotic cells; (2) an ampicillin resistance gene for selection of plasmid-containing prokaryotic cells; (3) an SV40 origin of replication for propagation in eukaryotic cells; (4) a CMV promoter, a polylinker, an SV40 intron; (5) several codons encoding a hemagglutinin fragment (i.e., an "HA" tag to facilitate purification) followed by a termination codon and polyadenylation signal arranged so that a DNA can be conveniently placed under expression control of the CMV promoter and operably linked to the SV40 intron and the polyadenylation signal by means of restriction sites in the polylinker. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein described by Wilson et al. 1984 Cell 37:767. The fusion of the HA tag to the target protein allows easy detection and recovery of the recombinant protein with an antibody that recognizes the HA epitope. pDNAIII contains, in addition, the selectable neomycin marker.

A DNA fragment encoding a *B. burgdorferi* polypeptide is cloned into the polylinker region of the vector so that recombinant protein expression is directed by the CMV promoter. The plasmid construction strategy is as follows. The DNA from a *B. burgdorferi* genomic DNA prep is amplified using primers that contain convenient restriction sites, much as described above for construction of vectors for expression of *B. burgdorferi* in *E. coli*. The 5' primer contains a Kozak sequence, an AUG start codon, and nucleotides of the 5' coding region of the *B. burgdorferi* polypeptide. The 3' primer, contains nucleotides complementary to the 3' coding sequence of the *B. burgdorferi* DNA, a stop codon, and a convenient restriction site.

The PCR amplified DNA fragment and the vector, pDNAI/Amp, are digested with appropriate restriction enzymes and then ligated. The ligation mixture is transformed into an appropriate *E. coli* strain such as SURE™ (Stratagene Cloning Systems, La Jolla, CA 92037), and the transformed culture is plated on ampicillin media plates which then are incubated to allow growth of ampicillin resistant colonies. Plasmid DNA is isolated from resistant colonies and examined by restriction analysis or other means for the presence of the fragment encoding the *B. burgdorferi* polypeptide.

For expression of a recombinant *B. burgdorferi* polypeptide, COS cells are transfected with an expression vector, as described above, using DEAE-dextran, as described, for instance, by Sambrook et al. (*supra*). Cells are incubated under conditions for expression of *B. burgdorferi* by the vector.

Expression of the *B. burgdorferi*-HA fusion protein is detected by radiolabeling and immunoprecipitation, using methods described in, for example Harlow et al., *supra*.. To this end, two days after transfection, the cells are labeled by incubation in media containing ³⁵S-cysteine for 8 hours. The cells and the media are collected, and the cells are washed and the lysed with detergent-containing RIPA buffer: 150 mM NaCl, 1% NP-40, 0.1% SDS, 1% NP-40, 0.5% DOC, 50 mM TRIS, pH 7.5, as described by Wilson et al. (*supra*). Proteins are

precipitated from the cell lysate and from the culture media using an HA-specific monoclonal antibody. The precipitated proteins then are analyzed by SDS-PAGE and autoradiography. An expression product of the expected size is seen in the cell lysate, which is not seen in negative controls.

5

8. Cloning and Expression in CHO Cells

The vector pC4 is used for the expression of *B. burgdorferi* polypeptide in this example. Plasmid pC4 is a derivative of the plasmid pSV2-dhfr (ATCC Accession No. 37146). The plasmid contains the mouse DHFR gene under control of the SV40 early promoter. Chinese hamster ovary cells or other cells lacking dihydrofolate activity that are transfected with these plasmids can be selected by growing the cells in a selective medium (alpha minus MEM, Life Technologies) supplemented with the chemotherapeutic agent methotrexate. The amplification of the DHFR genes in cells resistant to methotrexate (MTX) has been well documented. See, e.g., Alt et al., 1978, J. Biol. Chem. 253:1357-1370; Hamlin et al., 1990, Biochem. et Biophys. Acta, 1097:107-143; Page et al., 1991, Biotechnology 9:64-68. Cells grown in increasing concentrations of MTX develop resistance to the drug by overproducing the target enzyme, DHFR, as a result of amplification of the DHFR gene. If a second gene is linked to the DHFR gene, it is usually co-amplified and over-expressed. It is known in the art that this approach may be used to develop cell lines carrying more than 1,000 copies of the amplified gene(s). Subsequently, when the methotrexate is withdrawn, cell lines are obtained which contain the amplified gene integrated into one or more chromosome(s) of the host cell.

Plasmid pC4 contains the strong promoter of the long terminal repeat (LTR) of the Rouse Sarcoma Virus, for expressing a polypeptide of interest, Cullen, et al. (1985) Mol. Cell. Biol. 5:438-447; plus a fragment isolated from the enhancer of the immediate early gene of human cytomegalovirus (CMV), Boshart, et al., 1985, Cell 41:521-530. Downstream of the promoter are the following single restriction enzyme cleavage sites that allow the integration of the genes: *Bam* HI, *Xba* I, and *Asp* 718. Behind these cloning sites the plasmid contains the 3' intron and polyadenylation site of the rat preproinsulin gene. Other high efficiency promoters can also be used for the expression, e.g., the human β -actin promoter, the SV40 early or late promoters or the long terminal repeats from other retroviruses, e.g., HIV and HTLV. Clontech's Tet-Off and Tet-On gene expression systems and similar systems can be used to express the *B. burgdorferi* polypeptide in a regulated way in mammalian cells (Gossen et al., 1992, Proc. Natl. Acad. Sci. USA 89:5547-5551. For the polyadenylation of the mRNA other signals, e.g., from the human growth hormone or globin genes can be used as well. Stable cell lines carrying a gene of interest integrated into the chromosomes can also be selected upon co-transfection with a selectable marker such as gpt, G418 or hygromycin. It is advantageous to use more than one selectable marker in the beginning, e.g., G418 plus methotrexate.

The plasmid pC4 is digested with the restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from

a 1% agarose gel. The DNA sequence encoding the *B. burgdorferi* polypeptide is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the desired portion of the gene. A 5' primer containing a restriction site, a Kozak sequence, an AUG start codon, and nucleotides of the 5' coding region of the *B. burgdorferi* polypeptide is synthesized and used. A 3' primer, containing a restriction site, stop codon, and nucleotides complementary to the 3' coding sequence of the *B. burgdorferi* polypeptides is synthesized and used. The amplified fragment is digested with the restriction endonucleases and then purified again on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC4 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene are used for transfection. Five μ g of the expression plasmid pC4 is cotransfected with 0.5 μ g of the plasmid pSVneo using a lipid-mediated transfection agent such as Lipofectin™ or LipofectAMINE™ (Life Technologies Gaithersburg, MD). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μ M, 2 μ M, 5 μ M, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100-200 μ M. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

The disclosure of all publications (including patents, patent applications, journal articles, laboratory manuals, books, or other documents) cited herein are hereby incorporated by reference in their entireties.

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the invention. Functionally equivalent methods and components are within the scope of the invention, in addition to those shown and described herein and will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Such modifications are intended to fall within the scope of the appended claims.

TABLE 1. *Borrelia burgdorferi* - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident
1	92	100363	100184	gi1500722	similar to entire extracellular domain of glycine receptors [Caenorhabditis elegans]	100	66
1	537	513231	513608	gi147453	ribosomal protein S12 [Streptococcus pneumoniae]	92	85
1	283	272186	270849	gi1001376	ATP-dependent protease ATPase subunit [Synecocystis sp.]	89	75
1	847	798835	799131	gi1467373	ribosomal protein S18 [Bacillus subtilis]	86	69
1	78	91504	91235	gi1573896	ribosomal protein L27 (rpL27) [Haemophilus influenzae]	85	70
1	732	687538	686753	gi1591672	phosphate transport system ATP-binding protein [Methanococcus jannaschii]	84	65
1	788	739513	739232	gi142459	initiation factor 1 [Bacillus subtilis]	84	68
1	960	901448	901780	gnlPIDle2437	ORF YGL149w [Saccharomyces cerevisiae]	84	68
1	760	717009	715843	gi1623028	orf 361; translated orf similarity to SW: RF1_SALTY peptide chain release factor 1 of Salmonella typhimurium [Coxiella burnetii]	83	60
1	115	115536	115312	gi1695315	NADH dehydrogenase subunit [Digitalis grandiflora]	82	58
1	184	178954	176918	bbs157690	EF-G=elongation factor G [Thermotoga maritima, Peptide, 682 aa] [Thermotoga maritima]	82	63
1	447	425980	425453	gi1143804	Ndk [Bacillus subtilis]	82	56
1	201	194702	194103	gi1530438	arabinose transport protein [Mycoplasma capricolum]	81	53
1	477	446671	445589	gi1882454	fructose 1,6-bisphosphate aldolase [Escherichia coli]	81	61
1	601	569453	568650	gi1349227	transmembrane protein [Escherichia coli]	81	56
1	887	838084	837224	gi11237019	Srb [Bacillus subtilis]	81	52
1	889	840561	839497	gi1154276	peptide chain release factor 2 [Salmonella typhimurium]	81	65
1	896	846681	845440	gi1377823	aminopeptidase [Bacillus subtilis]	81	60
1	60	71604	68890	gi11619909	DNA mismatch repair protein [Thermotoga maritima]	80	59
1	354	348744	349157	gi11765976	chemotaxis protein CheY [Treponema pallidum]	80	42
1	423	409238	408855	gnlPIDle2118	50S ribosomal protein L14 [Odontella sinensis]	80	61
1	426	410130	409711	gi11652420	50S ribosomal protein L16 [Synecocystis sp.]	80	59
1	507	482736	482936	gi1515924	glucosyltransferase [Saccharomyces cerevisiae]	80	40
1	534	505081	505467	pirA02771IR	ribosomal protein L7/L12 - Micrococcus luteus	80	67

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

					7MCM			
1	597	567506	566532	gi580899	OppF gene product [Bacillus subtilis]			80
1	6	11241	9994	gnlPIDle2426 14	arginine deiminase [Clostridium perfringens]			79
1	478	447926	446835	gnlPIDle2881 24	glucose epimerase [Bacillus thuringiensis]			79
1	804	758549	757704	gi455176	glucosamine-6-phosphate deaminase protein [Escherichia coli]			79
1	25	31595	31894	gil1017809	similar to dihydropyridine-sensitive I-type, skeletal muscle calcium channel alpha-I subunit (SP:CIC1_RABIT, P07293) [Caenorhabditis elegans]			78
1	134	134667	134323	gil159199	cecropin D [Hyalophora cecropia]			78
1	230	215177	216028	gnlPIDle2655 37	DnaI-homologue [Thermus aquaticus thermophilus]			78
1	531	503406	503849	gi587583	ribosomal protein L11 [Thermus aquaticus thermophilus]			78
1	867	817849	819579	gi912449	Na+ -ATPase alpha subunit [Enterococcus hirae]			78
1	127	127383	127745	gi537364	heat shock protein 60 (GroEL) like protein [Porphyromonas gingivalis]			77
1	190	182991	182251	gil1235682	mevalonate pyrophosphate decarboxylase [Homo sapiens]			77
1	225	213158	212388	gil1651340	Phosphoglycerate mutase 1 [Escherichia coli]			77
1	284	272770	272165	gil1001349	ATP-dependent protease ClpP [Synecocystis sp.]			77
1	324	318280	314789	gil1573746	DNA polymerase III, alpha chain (dnaE) [Haemophilus influenzae]			77
1	555	530150	531370	gil143795	transfer RNA-Tyr synthetase [Bacillus subtilis]			77
1	770	722470	722892	gil1653602	hypothetical protein [Synecocystis sp.]			77
1	833	790115	790909	gnlPIDle2488 86	unknown [Mycobacterium tuberculosis]			77
1	52	62205	61918	gnlPIDle1189 66	ribosomal protein S15 [Thermus aquaticus thermophilus]			76
1	144	141975	141736	bbs177721	KHS toxin, killer heat sensitive toxin=KHS [Saccharomyces cerevisiae, Peptide, 708 aa] [Saccharomyces cerevisiae]			76
1	293	280702	280529	gil1146275	VP2 protein [Bluetongue virus 9]			76
1	323	314795	314199	gil1651915	hypothetical protein [Synecocystis sp.]			76
1	362	356749	355508	gil633147	ribose-phosphate pyrophosphokinase [Bacillus caldolyticus]			76

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	410	403332	402922	gil606232	30S ribosomal subunit protein S11 [Escherichia coli]	76	52
1	411	403754	403341	gil652405	30S ribosomal protein S13 [Synecocystis sp.]	76	55
1	454	431743	431003	gil1016012	neural cell adhesion protein BIG-2 precursor [Rattus norvegicus]	76	61
1	710	670457	671569	gil467376	unknown [Bacillus subtilis]	76	58
1	873	824849	826675	gil1303804	YqeQ [Bacillus subtilis]	76	52
1	942	886017	886751	gil183839	unknown [Pseudomonas aeruginosa]	76	54
1	5	9956	8943	gil1552842	OTCase [Escherichia coli]	75	62
1	51	61909	59735	gil1184680	polynucleotide phosphorylase [Bacillus subtilis]	75	54
1	55	66283	63620	gil39954	IF-2 (aa 1-741) [Bacillus stearothermophilus]	75	53
1	83	93454	94410	gnlPIDe289138	similar to flagellar hook-basal body proteins [Bacillus subtilis]	75	46
1	88	97435	98283	gil687583	RpoS [Yersinia enterocolitica]	75	47
1	245	229112	230158	gil1574806	spermidine/putrescine transport ATP-binding protein (potA) [Haemophilus influenzae]	75	55
1	264	251076	250801	gil1763634	alpha A-voltage-dependent calcium channel [Homo sapiens]	75	60
1	297	285723	284461	gil556886	serine hydroxymethyltransferase [Bacillus subtilis]	75	58
1	375	367682	366903	gil467372	3'-exo-deoxyribonuclease [Bacillus subtilis]	75	62
1	385	378055	377114	gil45986	NAD synthetase [Rhodobacter capsulatus]	75	55
1	416	406437	405925	gil1044981	ribosomal protein S5 [Bacillus subtilis]	75	56
1	418	407390	406812	gil600032	L6 ribosomal protein [Streptomyces coelicolor]	75	53
1	424	409520	409251	gil44218	ribosomal protein S17 (AA 1-85) [Mycoplasma capricolum]	75	58
1	530	502806	503366	gil396321	nusG [Escherichia coli]	75	56
1	548	523428	522904	gil1573470	H. influenzae predicted coding region HI0491 [Haemophilus influenzae]	75	55
1	575	546579	548393	pirIC300101C30010	hypothetical ORF-6 protein - Sauroleishmania tarentolae mitochondrion (SGC6)	75	50
1	906	854433	855215	gil511148	hemolysin [Serpulina hyodysenteriae]	75	56
1	68	85054	83102	gil467458	cell division protein [Bacillus subtilis]	74	57
1	162	158608	157502	gil531460	Mbl protein [Bacillus subtilis]	74	49
1	177	172327	171950	pirIA45434IA45434	ribosomal protein L19 - Bacillus stearothermophilus	74	54
1	475	443773	445203	gil396501	aspartyl-tRNA synthetase [Thermus aquaticus thermophilus]	74	52

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	549	524561	523374	gii1020317	S-adenosylmethionine synthetase [Staphylococcus aureus]	74	57
1	595	565672	564347	gii460259	enolase [Bacillus subtilis]	74	58
1	720	681529	680489	gii1651962	hypothetical protein [Synecocystis sp.]	74	49
1	745	702297	701173	gii289287	UDP-glucose pyrophosphorylase [Bacillus subtilis]	74	50
1	13	20409	17551	gii1652531	excinuclease ABC subunit A [Synecocystis sp.]	73	56
1	98	103790	104947	gii514330	sensor kinase [Bacillus subtilis]	73	49
1	188	182064	181102	gii887601	Erg8p [Saccharomyces cerevisiae]	73	43
1	314	303616	302786	gii473817	'ORF' [Escherichia coli]	73	53
1	366	358916	361078	gii1612457	ORF YLR069c [Saccharomyces cerevisiae]	73	51
			91				
1	444	424047	423181	gii1574704	hypothetical [Haemophilus influenzae]	73	51
1	556	531372	533672	gii511145	hemolysin [Serpulina hyodysenteriae]	73	52
1	576	548257	548045	gii406135	glycoprotein 120 [Simian immunodeficiency virus]	73	53
1	598	568379	567504	gii143607	sporulation protein [Bacillus subtilis]	73	55
1	604	572375	570729	bbs1161785	60 kDa antigen [Borrelia coriaceae, C053, ATCC 4338, Peptide, 514 aa] [Borrelia coriaceae]	73	53
1	674	634175	633648	gii1595810	type-I signal peptidase SpsB [Staphylococcus aureus]	73	47
1	692	654267	651727	gii1612684	unknown [Mycobacterium tuberculosis]	73	54
			56				
1	719	679186	680499	gii500705	Similar to Seryl-tRNA synthetase [Saccharomyces cerevisiae]	73	56
1	725	682189	682899	gii1612436	ORF YGR248w [Saccharomyces cerevisiae]	73	63
			81				
1	895	845455	844964	gii1652288	hypothetical protein [Synecocystis sp.]	73	50
1	16	24242	26497	gii511145	hemolysin [Serpulina hyodysenteriae]	72	53
1	99	104935	106305	gii619917	NtrC/NifA-like protein regulator [Escherichia coli]	72	54
1	133	134036	135055	gii556881	Similar to Saccharomyces cerevisiae SUA5 protein [Bacillus subtilis]	72	51
1	270	256925	260308	gii467444	transcription-repair coupling factor [Bacillus subtilis]	72	49
1	280	267529	268221	gii1573812	ribosomal protein S4 (rpS4) [Haemophilus influenzae]	72	51
1	282	270922	268472	gii402504	lon protease [Bacillus brevis]	72	51
1	325	319544	318363	gii48362	haemolysin releasing protein (AA 1-548) [Vibrio cholerae]	72	41
1	328	322678	321053	gii1591801	CTP synthase [Methanococcus jannaschii]	72	42

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	348	341460	341182	gill45687	ptsH protein [Escherichia coli]		72	55
1	405	399941	399096	gill1045937	M. genitalium predicted coding region MG246 [Mycoplasma genitalium]		72	53
1	420	408009	407779	gill580930	S14 protein (AA 1-61) [Bacillus subtilis]		72	55
1	593	563383	563850	gill574283	ribosomal protein L13 (rpL13) [Haemophilus influenzae]		72	54
1	682	641030	643399	gill574437	sporulation protein (spoIIIE) [Haemophilus influenzae]		72	51
1	754	710160	710750	gill460080	D-alanine:D-alanine ligase-related protein [Enterococcus faecalis]		72	47
1	767	721422	721640	gill868029	large ribosomal subunit protein L35 [Buchnera aphidicola]		72	48
1	860	811923	810511	gill1001357	asparaginyl-tRNA synthetase [Synecocystis sp.]		72	54
1	14	22434	20407	gill1737482	UvrB [Helicobacter pylori]		71	52
1	72	87471	87674	gill1016781	beta-b protein [Barley stripe mosaic virus]		71	42
1	289	278760	278239	gill534842	ORF9 [Rhizobium meliloti]		71	43
1	307	298685	296736	gill1652099	long-chain-fatty-acid CoA ligase [Synecocystis sp.]		71	48
1	321	313551	312130	gill1732243	RecG [Treponema pallidum]		71	52
1	522	494911	496383	gill459009	similar to multifunctional aminoacyl-tRNA synthetase, especially to the prolyl-tRNA synthetase region [Caenorhabditis elegans]		71	48
1	554	528795	530156	pirS58522IS58522	glycyl-tRNA synthetase - Thermus thermophilus		71	54
1	582	553725	552271	gill285623	pyruvate kinase [Bacillus stearothermophilus]		71	52
1	684	644626	643661	gill217121	ORF1 [Synecococcus elongatus]		71	52
1	723	681731	681561	gill44228	secretion protein SecY (AA 1-482) [Mycoplasma capricolum]		71	42
1	856	806939	807700	gill216341	ORF for methionine amino peptidase [Bacillus subtilis]		71	53
1	947	890096	890665	gill147485	queA [Escherichia coli]		71	56
1	28	38112	40613	gill1439562	Cdc28p [Schizosaccharomyces pombe]		70	53
1	36	45750	44806	gill290494	o287 [Escherichia coli]		70	32
1	84	94408	95220	gill47677	flgG protein product (AA 1-260) [Salmonella typhimurium]		70	50
1	128	127889	128569	gill1574387	H. influenzae predicted coding region HI1534 [Haemophilus influenzae]		70	58
1	468	441049	441330	gill1673757	(AE000012) Mycoplasma pneumoniae, phosphocarrier protein HPr; similar to GenBank Accession Number A49683, from M. capricolum [Mycoplasma pneumoniae]		70	41
1	532	503834	504529	spQ06797IRL1_BACSU	50S RIBOSOMAL PROTEIN L1 (BL1).		70	48

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	594	563858	564280	gil606169	30S ribosomal subunit protein S9 [Escherichia coli]	70	56
1	622	591070	591606	gil153906	CheW protein [Salmonella typhimurium]	70	48
1	703	664161	662611	gnlPIDle2839 19	glycerol kinase [Sulfolobus solfataricus]	70	60
1	726	682886	682659	gil836815	cdc4 gene product which is essential for initiation of DNA replication in yeast [Saccharomyces cerevisiae]	70	35
1	766	720854	721417	gil436165	Dsg [Myxococcus xanthus]	70	47
1	768	721649	722008	gnlPIDle2549 81	ribosomal protein L20 [Bacillus subtilis]	70	48
1	965	904395	905465	gil1100074	tryptophanyl-tRNA synthetase [Clostridium longisporum]	70	47
1	87	96986	97336	gil160092	asparagine-rich protein [Plasmodium falciparum]	69	46
1	110	112658	113602	gil1001733	ABC transporter [Synechocystis sp.]	69	46
1	181	174037	173762	pirC47154IC 47154	ribosomal protein S16 - Bacillus subtilis	69	52
1	233	219872	218076	gil1001493	protein-export membrane protein SecD [Synechocystis sp.]	69	47
1	234	220245	219922	gil1402532	ORF11 [Enterococcus faecalis]	69	32
1	373	366148	363977	gil1574200	hypothetical [Haemophilus influenzae]	69	48
1	419	407781	407371	gil498771	ribosomal S8 protein [Thermus aquaticus thermophilus]	69	46
1	517	489315	491207	gil151932	fructose enzyme II [Rhodobacter capsulatus]	69	42
1	600	568891	568388	gil143606	sporulation protein [Bacillus subtilis]	69	44
1	733	689098	687536	gil1303856	YggI [Bacillus subtilis]	69	46
1	874	826778	827746	pirS08183IS0 8183	L-lactate dehydrogenase (EC 1.1.1.27) X - Bacillus psychrosaccharolyticus	69	50
1	894	844392	844547	gil1592324	M. jannaschii predicted coding region MJ1172 [Methanococcus jannaschii]	69	53
1	934	879725	879237	gil153566	ORF (19K protein) [Enterococcus faecalis]	69	42
1	49	57779	57976	gil809583	unknown [Saccharomyces cerevisiae]	68	36
1	107	110374	111513	gnlPIDle2559 43	M04B2.4 [Caenorhabditis elegans]	68	48
1	132	133978	133148	gil1001663	rare lipoprotein A [Synechocystis sp.]	68	53
1	142	141239	142642	gnlPIDle2338 74	hypothetical protein [Bacillus subtilis]	68	45
1	148	145381	144005	gil558574	pyrophosphate-fructose-6-phosphate 1-phosphotransferase	68	48

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

[illegible]

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	111	113572	114333	gil1001529	hypothetical protein [Synecocystis sp.]	67	36
1	170	166286	165876	gil567036	CapE [Staphylococcus aureus]	67	35
1	202	195499	194651	gil1674275	(AE000056) Mycoplasma pneumoniae, hypothetical ABC transporter (yicW) homolog; similar to Swiss-Prot Accession Number P32721, from E. coli [Mycoplasma pneumoniae]	67	41
1	206	197487	197098	gil1653841	P protein [Synecocystis sp.]	67	35
1	271	260292	261551	gil349834	acetate kinase [Methanosarcina thermophila]	67	44
1	313	302731	301643	gnlPIDle249981	phosphotransacetylase [Thermoanaerobacterium thermosaccharolyticum]	67	51
1	422	408897	408535	pirA02819IR5BS24	ribosomal protein L24 - Bacillus stearothermophilus	67	49
1	480	450326	448689	gil1574032	hypothetical [Haemophilus influenzae]	67	42
1	529	502315	502509	gil1001264	50S ribosomal protein L33 [Synecocystis sp.]	67	56
1	588	559618	561111	gil1224069	amidase [Moraxella catarrhalis]	67	51
1	683	643676	643437	gil710340	ribosomal protein S21 [Myxococcus xanthus]	67	49
1	698	658454	659500	gil460955	TagE [Vibrio cholerae]	67	38
1	700	660039	660536	gil467420	unknown [Bacillus subtilis]	67	42
1	729	684089	685888	gnlPIDle267607	alanyl-tRNA synthetase [Thermus aquaticus thermophilus]	67	51
1	835	791754	792341	gnlPIDle248763	unknown [Mycobacterium tuberculosis]	67	46
1	857	807722	809191	gil1526428	GsrA protein [Yersinia enterocolitica]	67	46
1	868	819577	820905	gil1590954	ATP synthase, subunit B [Methanococcus jannaschii]	67	53
1	74	88393	88028	gil1572979	hypothetical [Haemophilus influenzae]	66	43
1	91	99152	100252	gil561690	sialoglycoprotease [Pasteurella haemolytica]	66	44
1	123	121472	120783	gil1652843	endonuclease III [Synecocystis sp.]	66	42
1	149	146362	145379	gil1216385	orf304 gene product [Treponema pallidum]	66	43
1	185	179585	179001	gil1574811	neutrophil activating protein (napA) [Haemophilus influenzae]	66	49
1	275	265075	265584	gil401785	cytidine deaminase [Mycoplasma pirum]	66	41
1	330	324514	323696	gil1574641	ribonucleotide transport ATP-binding protein (mkl) [Haemophilus influenzae]	66	41
1	335	327265	326888	gil510670	cheY gene product [Rhodobacter sphaeroides]	66	44
1	355	349142	349603	gil499382	FliS [Bacillus subtilis]	66	28

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	358	351051	350827	gil1546788	tar-1 [Trichostrongylus colubriformis]	66	55
1	404	399121	398324	gil296626	hemolysin [Serpulina hyodysenteriae]	66	53
1	491	461335	460550	gil45713	P.putida genes rpmH, rnpA, 9k, 60k, 50k, gidA, gidB, uncl and uncB [Pseudomonas putida]	66	41
1	513	486046	485159	gil153903	methyltransferase (cheR; EC 2.1.1.24) [Salmonella typhimurium]	66	42
1	552	526495	527316	gil340613	A 'c' was inserted after nt 369 (=nt 10459 in genomic sequence (M10126)) to correct -1 frameshift probably due to gel compression [Leishmania tarentolae]	66	40
1	611	579933	581069	gil886130	putative pectinesterase [Medicago sativa]	66	33
1	627	595395	596288	gilPIDle2639	OrfD [Streptococcus pneumoniae]	66	47
1	772	723788	723522	gil1762342	could accelerate degradation of certain transcripts [Bacillus subtilis]	66	47
1	816	770251	770060	gil393266	glycerol ester hydrolase [Staphylococcus aureus]	66	33
1	841	795927	795208	gil662880	novel hemolytic factor [Bacillus cereus]	66	46
1	882	835002	834262	gil862629	similar to the ATP-binding transport protein family [Buchnera aphidicola]	66	40
1	73	87915	87619	gil39656	spoVG gene product [Bacillus megaterium]	65	40
1	97	103039	102803	gil532272	phosphatidylserine decarboxylase [Bacillus subtilis]	65	39
1	106	110281	109649	gil377852	ClpP [Yersinia enterocolitica]	65	42
1	159	156186	154372	gil572977	penicillin-binding protein 2 (pbp2) [Haemophilus influenzae]	65	41
1	172	168084	169325	gil146238	poly(A) polymerase [Bacillus subtilis]	65	38
1	268	255918	253819	gil829194	bacterial cell wall hydrolase [Enterococcus faecalis]	65	43
1	353	348568	346553	gil574651	DNA ligase (lig) [Haemophilus influenzae]	65	45
1	696	657577	655781	gil651216	Pz-peptidase [Bacillus licheniformis]	65	47
1	741	695297	693456	gil575784	DNA mismatch repair protein [Aquifex pyrophilus]	65	45
1	846	798339	798827	gil1001362	single-stranded DNA-binding protein [Synecocystis sp.]	65	45
1	932	876643	878559	gil508471	gyrase A [Helicobacter pylori]	65	40
1	936	881238	882224	gil652260	leader peptidase I [Synecocystis sp.]	65	40
1	961	902331	901519	gil1256146	YbbQ [Bacillus subtilis]	65	48
1	963	903280	904407	gil1573307	hypothetical [Haemophilus influenzae]	65	41
1	37	47101	45683	gil556014	UDP-N-acetyl muramate-alanine ligase [Bacillus subtilis]	64	46

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	61	72211	71642	gill041785	rhostry protein [Plasmodium yoelii]	64	41
1	130	131969	129336	gill574225	valyl-tRNA synthetase (valS) [Haemophilus influenzae]	64	43
1	156	152924	151140	gill43066	threonyl-tRNA synthetase (thrS; EC 6.1.1.3) [Escherichia coli]	64	43
1	174	170326	170033	gill652390	acyl carrier protein [Synecocystis sp.]	64	48
1	175	171105	170545	gill573650	lipopolysaccharide core biosynthesis protein (kdtB) [Haemophilus influenzae]	64	45
1	178	173033	172293	gill1046163	tRNA (guanine-N1)-methyltransferase [Mycoplasma genitalium]	64	47
1	180	173764	173513	gnlPIDle248893	unknown [Mycobacterium tuberculosis]	64	34
1	207	197654	197436	gill11665	ORF2136 [Marchantia polymorpha]	64	47
1	217	206795	205761	gill652866	N-acetylmuramoyl-L-alanine amidase [Synecocystis sp.]	64	30
1	244	228146	229036	gill1046160	hypothetical protein (GB:U00021_5) [Mycoplasma genitalium]	64	33
1	246	230149	230967	gill147336	transmembrane protein [Escherichia coli]	64	36
1	267	253160	253723	gill467430	unknown [Bacillus subtilis]	64	52
1	340	333349	332783	gill145520	cheW peptide [Escherichia coli]	64	42
1	384	376509	375565	gill653737	monophosphatase [Synecocystis sp.]	64	44
1	449	428137	426437	gill467409	DNA polymerase III subunit [Bacillus subtilis]	64	41
1	510	484558	483998	pirA005471X YEBET	protein-glutamate methyltransferase (EC 3.1.1.61) - Salmonella typhimurium	64	45
1	603	570416	569451	gill1574678	dipeptide transport system permease protein (dppB) [Haemophilus influenzae]	64	44
1	679	637996	640224	gill1001335	soluble lytic transglycosylase [Synecocystis sp.]	64	42
1	753	709637	710194	gnlPIDle283360	unknown [Mycobacterium tuberculosis]	64	42
1	817	771784	771969	gnlPIDle250307	W04B2.3 [Caenorhabditis elegans]	64	41
1	839	793892	795211	gill1573939	hypothetical [Haemophilus influenzae]	64	38
1	861	811972	812853	gill396314	glutamate synthase [Escherichia coli]	64	38
1	870	821501	823339	gill472918	v-type Na-ATPase [Enterococcus hirae]	64	42
1	901	850668	851615	gill581088	methionyl-tRNA formyltransferase [Escherichia coli]	64	38
1	904	853492	853884	gill992960	thioredoxin [Arabidopsis thaliana]	64	41
1	24	34314	31444	gill42914	SbcC (AA 1-1048) [Escherichia coli]	63	45

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	77	91198	90194	gil1652022	GTP-binding protein [Synecocystis sp.]	63	45
1	209	198041	197862	gil11665	ORF2136 [Marchantia polymorpha]	63	42
1	227	214639	213956	gil1652349	oxygen independent coprophorphyrinogen III oxidase [Synecocystis sp.]	63	41
1	232	218116	217193	gil1573204	protein-export membrane protein (secF) [Haemophilus influenzae]	63	36
1	247	230965	231762	sp145169IPO	SPERMIDINE/PUTRESCINE TRANSPORT SYSTEM	63	36
1	272	262171	261614	TC_HAEIN	PERMEASE PROTEIN POTC.		
1	290	279964	278735	gil1780755	DJ-1 protein [Homo sapiens]	63	46
1	333	326012	325818	gil143439	DD-carboxypeptidase [Bacillus subtilis]	63	41
1	508	484000	482759	gil293954	mating type a-1 protein [Neurospora crassa]	63	42
1	553	527314	528801	gil1041116	TRAB [Plasmid pPD1]	63	34
1	569	542317	543747	sp15189ISY	GLUTAMYL-TRNA SYNTHETASE (EC 6.1.1.17)	63	48
1	620	590442	589360	E_RHIME	(GLUTAMATE-TRNA LIGASE) (GLURS).		
1	701	660784	660623	gil1652577	carboxyl-terminal protease [Synecocystis sp.]	63	49
1	702	662231	660735	gil1098641	Bts1p [Saccharomyces cerevisiae]	63	43
1	704	664938	664159	gil1339938	EC 1.1.99.5 [Mus musculus]	63	47
1	746	702035	702631	gil763191	glycerol 3 phosphate dehydrogenase [Saccharomyces cerevisiae]	63	37
1	748	705645	704671	gil142997	glycerol uptake facilitator [Bacillus subtilis]	63	45
1	749	706431	705643	gil41497	ORF 4 (AA 1-198); 20 kD [Escherichia coli]	63	36
1	756	715040	713019	gil436158	putative integral membrane protease required for high frequency lysogenization by bacteriophage lambda [Escherichia coli]	63	33
1	825	780572	783289	gil507734	HflK [Vibrio parahaemolyticus]	63	28
1	853	803786	804832	gil407881	stringent response-like protein [Streptococcus equisimilis]	63	40
1	4	8945	7467	gil746399	transcription elongation factor [Escherichia coli]	63	43
1	42	50587	51786	gil155055	basic membrane protein precursor [Treponema pallidum]	63	36
1	57	67740	66271	gil1573583	H. influenzae predicted coding region HI0594 [Haemophilus influenzae]	62	38
1	64	78979	79767	gil149316	pantothenate metabolism flavoprotein (dfr) [Haemophilus influenzae]	62	41
1	80	92123	91806	gil1001473	ORF2 gene product [Bacillus subtilis]	62	36
1	80	92123	91806	gn11PID1e2118	hypothetical protein [Synecocystis sp.]	62	48
1	80	92123	91806	gn11PID1e2118	50S ribosomal protein L21 [Odontella sinensis]	62	34

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	101	107458	106793	gil1652679	hypothetical protein [Synechocystis sp.]	62	34
1	102	107464	107883	gil15893	predicted 12.5Kd protein [Mycobacteriophage 15]	62	44
1	226	213238	213969	gil1001678	ribose 5-phosphate isomerase [Synechocystis sp.]	62	39
1	266	251889	253175	gil529118	similar to APE1/LAP4, vacuolar aminopeptidase [Saccharomyces cerevisiae]	62	42
1	299	288749	287274	gil289284	cysteineyl-tRNA synthetase [Bacillus subtilis]	62	46
1	357	349982	352714	gil1633576	similar to proofreading 3'-5' exonuclease and polymerase [Treponema pallidum]	62	41
1	443	423190	422495	gil312380	putative orfW gene product [Clostridium acetobutylicum]	62	32
1	489	458740	459582	gil40031	spoJ93 gene product [Bacillus subtilis]	62	36
1	511	485147	484494	gil145524	cheB peptide [Escherichia coli]	62	28
1	518	491201	492322	gil146722	phosphomannose isomerase [Escherichia coli]	62	45
1	685	646727	644598	gil1574144	single-stranded-DNA-specific exonuclease (recJ) [Haemophilus influenzae]	62	40
1	695	655800	655063	gil1477770	unknown [Helicobacter pylori]	62	37
1	758	715668	714979	gil1574130	protoporphyrinogen oxidase (hemK) [Haemophilus influenzae]	62	36
1	762	718374	719198	gil1652444	hypothetical protein [Synechocystis sp.]	62	40
1	837	792941	793891	gil1652668	phosphatidate cytidyltransferase [Synechocystis sp.]	62	41
1	917	862498	862737	gil440851	collagenase [Clostridium perfringens]	62	29
1	46	55889	54726	gil498141	tRNA guanine transglycosylase [Zymomonas mobilis]	61	35
1	81	92710	92174	gil726305	adenine phosphoribosyltransferase form 1 [Triticum aestivum]	61	45
1	100	106820	106557	gil460955	TagE [Vibrio cholerae]	61	45
1	109	111699	112664	gil1001126	hypothetical protein [Synechocystis sp.]	61	48
1	157	154445	153051	gil143657	endospore forming protein [Bacillus subtilis]	61	40
1	193	185315	184227	gil48409	gene not found in Erwinia uredovora crt gene cluster; ORF6 [Erwinia herbicola]	61	42
1	223	209790	210668	spIP37214IER A_STRMU	GTP-BINDING PROTEIN ERA HOMOLOG.	61	37
1	273	262392	264062	gil438455	possible N-terminal signal sequence; mature protein may be membrane-anchored and start at Cys-17. 17.5% identity over 354-aa overlap with Candida pelliculosa beta-glucosidase.; putative [Bacillus subtilis]	61	37
1	277	265982	265581	gil1513240	ORFveg110 [Dictyostelium discoideum]	61	29

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	301	291935	289686	gill354776	MCP-1 [Treponema pallidum]	61	43
1	322	314201	313338	gill732243	RecG [Treponema pallidum]	61	37
1	380	371430	372392	gill973332	OrfC [Bacillus subtilis]	61	38
1	408	401874	401479	gill147716	ribosomal protein L17 [Escherichia coli]	61	44
1	413	404277	404444	gill185286	ORF [Sulfolobus shibatae]	61	47
1	415	405927	405616	pirA02827IR 5BS3F	ribosomal protein L30 - Bacillus stearothermophilus	61	31
1	417	406848	406435	pirB29102IR 5BS8F	ribosomal protein L18 - Bacillus stearothermophilus	61	44
1	441	421784	421224	gill153045	prolipoprotein signal peptidase [Staphylococcus aureus]	61	29
1	467	440722	441042	gill173128	ubiquitin-specific processing protease [Saccharomyces cerevisiae]	61	32
1	613	582695	581547	gill1303756	YqbP [Bacillus subtilis]	61	38
1	615	584397	585476	gill551522	TpN38(b) [Treponema pallidum]	61	26
1	673	632123	633622	gill143999	dnaK homologue [Borrelia burgdorferi]	61	41
1	675	634207	635469	gill1653709	lipoprotein NlpD [Synecocystis sp.]	61	50
1	743	699438	698647	gill1303863	YqgP [Bacillus subtilis]	61	45
1	897	847575	846688	gill1573586	hydrolase (GB:Z33006_1) [Haemophilus influenzae]	61	43
1	938	882836	883282	gill1303831	YqfM [Bacillus subtilis]	61	36
1	7	10415	10627	gnlPIDle2119 90	T24A11.1 [Caenorhabditis elegans]	60	45
1	23	31428	30475	gill1303865	YqgR [Bacillus subtilis]	60	45
1	35	44812	44267	gill1591369	cytidylate kinase [Methanococcus jannaschii]	60	49
1	198	192994	192053	gill1045801	hypothetical protein (SP:P32720) [Mycoplasma genitalium]	60	33
1	347	341167	339440	gill602680	phosphocarrier protein (enzyme I) [Mycoplasma capricolum]	60	37
1	369	361817	362233	gill1372995	OrfH [Borrelia burgdorferi]	60	37
1	409	402924	401872	gill142463	RNA polymerase alpha-core-subunit [Bacillus subtilis]	60	40
1	438	420142	418793	gnlPIDle2768 30	UDP-N-acetylglucosamine 1-carboxyvinyltransferase [Bacillus subtilis]	60	40
1	566	540696	539698	gill1573923	prolipoprotein diacylglycerol transferase (lgt) [Haemophilus influenzae]	60	57
1	587	559368	559655	gill1335805	CD45 homolog [Heterodontus francisci]	60	26
1	589	561098	562558	gill1653395	PET112 [Synecocystis sp.]	60	37

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	736	690957	690076	gill1001260	hypothetical protein [Synechocystis sp.]	60	47
1	738	691078	691659	gill1399829	elongation factor P [Synechococcus PCC7942]	60	34
1	750	707879	706626	gill1573060	hypothetical [Haemophilus influenzae]	60	33
1	784	734589	735635	gill1164996	mxnC gene product [Methylobacterium extorquens]	60	26
1	829	785899	786567	gill1046033	cytidylate kinase [Mycoplasma genitalium]	60	38
1	862	812835	813773	gill1574569	hypothetical [Haemophilus influenzae]	60	36
1	863	813727	816105	gill1046033	hypothetical protein [Bacillus subtilis]	60	38
1	878	831250	829943	gill1742766	NitS protein. [Escherichia coli]	60	34
1	929	872578	874110	gill1002666	unknown [Schistosoma mansoni]	60	30
1	937	882211	882861	gill1595810	type-I signal peptidase SpsB [Staphylococcus aureus]	60	40
1	54	63629	63234	gill580902	ORF6 gene product [Bacillus subtilis]	59	38
1	96	102744	103802	gill467409	DNA polymerase III subunit [Bacillus subtilis]	59	40
1	120	118925	119914	gill1574678	dipeptide transport system permease protein (dppB) [Haemophilus influenzae]	59	33
1	140	139567	141174	gill42377	phosphoglucose isomerase (AA 1-549) [Escherichia coli]	59	42
1	195	186577	187659	gill1573129	hypothetical [Haemophilus influenzae]	59	41
1	259	242174	245713	gill1574781	exodeoxyribonuclease V (recB) [Haemophilus influenzae]	59	38
1	288	278281	276257	pirD64084ID 64084	rep helicase, single-stranded DNA-dependent ATPase (rep) homolog - Haemophilus influenzae (strain Rd KW20)	59	36
1	291	280005	281525	gill882504	ORF_f560 [Escherichia coli]	59	34
1	306	294923	296707	gill487937	Similar to arginyl-tRNA synthetase (E. coli) [Saccharomyces cerevisiae]	59	35
1	332	325664	324564	gill466753	alternate gene name yibD [Escherichia coli]	59	39
1	414	405646	405179	gill216338	ORF for L15 ribosomal protein [Bacillus subtilis]	59	40
1	465	439470	440759	gill39269	sigma factor (ntrA) (AA 1-502) [Azotobacter vinelandii]	59	35
1	492	462064	461411	pirA30191IA 30191	hypothetical protein L - Bacillus subtilis (fragment)	59	39
1	495	462955	463752	gill467425	unknown [Bacillus subtilis]	59	38
1	503	480078	481016	gill1651878	regulatory components of sensory transduction system [Synechocystis sp.]	59	38
1	523	497621	496395	gill143002	proton glutamate symport protein [Bacillus caldotenax]	59	34

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	941	885060	886019	gil1685110	tetrahydrofolate dehydrogenase/cyclohydrolase [Streptococcus thermophilus]	59	36
1	40	50348	48951	gil1574003	pantothenate permease (panF) [Haemophilus influenzae]	58	38
1	76	90160	89534	gil1303791	YqeJ [Bacillus subtilis]	58	32
1	116	115845	115654	gnlPIDle275892	T06E6.f [Caenorhabditis elegans]	58	37
1	179	173515	173009	gil1573163	hypothetical [Haemophilus influenzae]	58	37
1	197	191904	189634	gil1066850	putative [Rhodobacter capsulatus]	58	37
1	229	215111	214563	gil1573441	oxygen-independent coproporphyrinogen III oxidase (hemN) [Haemophilus influenzae]	58	34
1	257	238952	241873	gil1041785	rhopty protein [Plasmodium yoelii]	58	30
1	440	421010	420792	gil1674178	(AE000047) Mycoplasma pneumoniae, MG246 homolog, from M. genitalium [Mycoplasma pneumoniae]	58	37
1	557	533653	534750	gil974332	NAD(P)H-dependent dihydroxyacetone-phosphate reductase [Bacillus subtilis]	58	41
1	586	557259	559370	gil153062	helicase [Staphylococcus aureus]	58	41
1	623	591542	592435	gil1653618	hypothetical protein [Synecocystis sp.]	58	35
1	728	683208	684104	gil790935	flgG [Treponema denticola]	58	31
1	796	750629	749508	gil1574412	alanine racemase, biosynthetic (alr) [Haemophilus influenzae]	58	29
1	823	778475	778723	gil1209836	minus strand repeat motif-containing gene [Borrelia burgdorferi]	58	22
1	830	786540	788225	gil1574150	ribosomal protein S1 (rpS1) [Haemophilus influenzae]	58	34
1	842	796255	796019	gnlPIDle243474	ORF YGR089w [Saccharomyces cerevisiae]	58	35
1	883	834332	834520	gil1575792	low Mr GTP-binding protein Rab32 [Homo sapiens]	58	43
1	905	853953	854435	gil1303823	YqfG [Bacillus subtilis]	58	34
1	919	863594	862875	gil1256625	putative [Bacillus subtilis]	58	34
1	921	865297	864725	gil1054584	putative protein highly homologous to E. coli RNase HII [Magnetospirillum sp.]	58	42
1	196	189636	187702	gil496484	tlpC gene product [Bacillus subtilis]	57	32
1	262	249142	248192	gil46605	lacC polypeptide (AA 1-310) [Staphylococcus aureus]	57	41
1	311	300776	301660	gil467431	high level kasamycin resistance [Bacillus subtilis]	57	35
1	365	358725	358495	gil396943	early protein [Human papillomavirus type 19]	57	38
1	386	378249	378025	gil45986	NAD synthetase [Rhodobacter capsulatus]	57	32

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	399	394690	394247	gil1592085	M. jannaschii predicted coding region MJ1437 [Methanococcus jannaschii]	57	31
1	402	397512	396193	gil1732241	GTP-binding protein [Treponema pallidum]	57	36
1	533	504504	505022	gil786163	Ribosomal Protein L10 [Bacillus subtilis]	57	29
1	735	689992	689096	gil1303855	YqgH [Bacillus subtilis]	57	30
1	794	745857	747644	gil11665	ORF2136 [Marchantia polymorpha]	57	33
1	814	768735	771866	gil1573914	acridine resistance protein (acrB) [Haemophilus influenzae]	57	31
1	821	776835	778244	gil1591660	histidyl-tRNA synthetase [Methanococcus jannaschii]	57	36
1	834	790907	791752	gil1518661	elongation factor Ts [Chlamydia trachomatis]	57	36
1	836	792328	793038	gil1573941	hypothetical [Haemophilus influenzae]	57	36
1	848	799086	799670	gil1537044	50S ribosomal subunit protein L9 [Escherichia coli]	57	39
1	849	799668	801041	gil1001271	replicative DNA helicase [Synechocystis sp.]	57	33
1	851	802510	803742	pirA64092IA 64092	acetyl coenzyme A acetyltransferase (thiolase) (fadA) homolog - Haemophilus influenzae (strain Rd KW20)	57	43
1	855	805240	806952	gil1499620	M. jannaschii predicted coding region MJ0798 [Methanococcus jannaschii]	57	36
1	922	865347	867809	gil1237015	ORF4 [Bacillus subtilis]	57	38
1	12	17611	14156	pirB30565IB 30565	phospholipase C (EC 3.1.4.3) precursor - Clostridium bifermentans	56	28
1	26	35530	34277	gil1657594	exonuclease SbcD [Escherichia coli]	56	39
1	59	68915	68271	gil148876	probable com101A gene [Haemophilus influenzae]	56	26
1	79	91821	91480	gil1139633	large tegument protein [Human herpesvirus 7]	56	30
1	112	113768	113571	gnlPIDle2469 33	ORF YPL216w [Saccharomyces cerevisiae]	56	35
1	147	142606	143988	gil1642030	NADH oxidase [Serpulina hyodysenteriae]	56	36
1	153	148561	149100	gil1499018	M. jannaschii predicted coding region MJ0240 [Methanococcus jannaschii]	56	29
1	169	165431	164388	gil1573431	aminodeoxychorismate lyase (pabC) [Haemophilus influenzae]	56	41
1	183	176655	175432	gil143841	xylose repressor [Bacillus subtilis]	56	38
1	312	301170	300922	gnlPIDle2202 40	red algal chloroplast [Plasmodium falciparum]	56	36
1	317	308362	306992	gil1574691	UDP-N-acetylmuramoylalanine-D-glutamate ligase (murD) [Haemophilus influenzae]	56	40

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	448	426477	426133	gil467410	unknown [Bacillus subtilis]	56	28
1	456	432628	434457	gil142521	deoxyribodipyrimidine photolyase [Bacillus subtilis]	56	34
1	460	438178	437312	gil882453	ORF_f286; alternate name yggB; orf4 of X14436 [Escherichia coli]	56	31
1	469	441309	443438	gil148316	NaH-antiporter protein [Enterococcus hirae]	56	32
1	608	574772	574951	gil1019630	NADH dehydrogenase subunit 2 [Paramecium aurelia]	56	37
1	699	659498	660055	gil1372995	OrfH [Borrelia burgdorferi]	56	24
1	757	713509	713712	gil861327	F31D5.5 gene product [Caenorhabditis elegans]	56	40
1	791	741305	742837	gil1651873	4-alpha-glucanotransferase [Synechocystis sp.]	56	43
1	822	779478	778291	gil1500309	M. jannaschii predicted coding region MJ1428 [Methanococcus jannaschii]	56	28
1	967	907556	908932	gil1749528	similar to Saccharomyces cerevisiae probable UTP-glucose-1-phosphate uridylyltransferase, SWISS-PROT Accession Number P32861 [Schizosaccharomyces pombe]	56	37
1	39	48953	48048	gil1045895	hypothetical protein (SP:P23851) [Mycoplasma genitalium]	55	41
1	131	132989	131967	gil1574007	nitrogen fixation nifR3 protein (nifR3) (PIR:S49971) [Haemophilus influenzae]	55	39
1	152	148506	147148	gil1653100	Na+ -ATPase subunit J [Synechocystis sp.]	55	31
1	359	352690	353313	gil1213334	OrfX; hypothetical 22.5 KD protein downstream of type IV prepilin leader peptidase gene; Method: conceptual translation supplied by author [Vibrio vulnificus]	55	33
1	361	355510	354140	gil882698	L-fuculose kinase [Escherichia coli]	55	44
1	515	488398	487652	gil397486	endonuclease G [Bos taurus]	55	33
1	551	526427	525285	gil558266	orf gene product [Wolinnella succinogenes]	55	30
1	570	543745	544482	gil1303811	YqeU [Bacillus subtilis]	55	33
1	579	551201	551494	gil290487	50S ribosomal subunit protein L28 [Escherichia coli]	55	37
1	584	555359	556063	gil1592301	M. jannaschii predicted coding region MJ0687 [Methanococcus jannaschii]	55	32
1	706	665310	665936	gil403984	deoxyguanosine kinase/deoxyadenosine kinase(I) subunit [Lactobacillus acidophilus]	55	38
1	771	722876	723538	gil1736440	O-sialoglycoprotein endopeptidase (EC 3.4.24.57) (Glycoprotease). [Escherichia coli]	55	39
1	786	736537	737187	gil1589778	SPINDLY [Arabidopsis thaliana]	55	34

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	810	765243	766130	gil984805	glycine betaine-binding protein precursor [Bacillus subtilis]	55	35
1	871	823341	823790	gil1590959	ATP synthase, subunit K [Methanococcus jannaschii]	55	34
1	898	847660	849462	gil1517942	aminopeptidase P [Sus scrofa]	55	46
1	924	867811	868236	gil1142660	POM1 [Plasmodium chabaudi chabaudi]	55	41
1	927	870905	870039	gil534839	CheR [Rhizobium meliloti]	55	32
1	964	904091	903900	gil312694	ARS-binding factor 1 [Kluyveromyces marxianus]	55	50
1	33	44068	43124	gil146860	delta-2-isopentenyl pyrophosphate transferase [Escherichia coli]	54	31
1	63	79094	74679	gil415736	Orf635 gene product [Euglena gracilis]	54	37
1	192	184282	182969	gil151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii]	54	35
1	200	194105	192951	gil1045800	ribose transport system permease protein [Mycoplasma genitalium]	54	29
1	224	210749	212320	gil1591243	M. jannaschii predicted coding region MJ0539 [Methanococcus jannaschii]	54	45
1	256	237491	238954	gnlPIDle245024	unknown [Mycobacterium tuberculosis]	54	34
1	260	245698	247542	gil1574782	exodeoxyribonuclease V (recD) [Haemophilus influenzae]	54	36
1	320	311333	312133	gil1209528	D,D-carboxypeptidase [Enterococcus faecalis]	54	40
1	610	577096	579909	gil1499043	M. jannaschii predicted coding region MJ0263 [Methanococcus jannaschii]	54	30
1	765	720685	719999	gil290216	[bride of sevenless] gene product [Drosophila virilis]	54	25
1	789	739607	739996	gil473804	'dosage-dependent dnaK suppressor protein' [Escherichia coli]	54	35
1	845	797932	798366	gil1045767	ribosomal protein S6 [Mycoplasma genitalium]	54	35
1	951	894898	893912	gil1303842	YqfU [Bacillus subtilis]	54	28
1	86	96019	97032	gil405550	flagellar P-ring protein [Pseudomonas putida]	53	40
1	89	98331	99215	gil912478	No definition line found [Escherichia coli]	53	35
1	164	159533	158562	gil1499620	M. jannaschii predicted coding region MJ0798 [Methanococcus jannaschii]	53	39
1	250	234276	232861	gil1303989	YqkI [Bacillus subtilis]	53	28
1	278	266053	267426	gil1749686	similar to Saccharomyces cerevisiae unknown, EMBL Accession Number Z68194 [Schizosaccharomyces pombe]	53	28
1	302	292150	294309	gil1015945	methyl accepting chemotaxis homolog [Treponema denticola]	53	31
1	364	358298	357702	gil1499620	M. jannaschii predicted coding region MJ0798 [Methanococcus jannaschii]	53	41

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	514	486253	486888	gil940842	orf 06111 gene product [Saccharomyces cerevisiae]	53	28
1	567	541832	540684	gil1165254	YlxH [Borrelia burgdorferi]	53	33
1	621	590418	591032	gil1592021	cell division protein J [Methanococcus jannaschii]	53	32
1	805	759748	758537	gil1732203	GlcNAc 6-P deacetylase [Vibrio fischeri]	53	35
1	854	804825	805298	gil1303915	YqbZ [Bacillus subtilis]	53	30
1	884	835705	834944	gil1574399	H. influenzae predicted coding region HI1555 [Haemophilus influenzae]	53	26
1	48	58236	56944	gil1652686	hypothetical protein [Synechocystis sp.]	52	26
1	53	63264	62383	gil42219	P35 gene product (AA 1 - 314) [Escherichia coli]	52	34
1	56	66168	65665	gil1151158	repeat organellar protein [Plasmodium falciparum]	52	37
1	95	102255	102746	gil1574136	colicin V production protein (pur regulon) (cvpA) [Haemophilus influenzae]	52	24
1	117	115800	116879	gil288998	secA gene product [Antithamnion sp.]	52	34
1	220	208898	208446	gil1652602	hypothetical protein [Synechocystis sp.]	52	25
1	285	274152	272764	gil1652602	trigger factor [Bacillus subtilis]	52	28
1	352	344946	346532	gil160299	glutamic acid-rich protein [Plasmodium falciparum]	52	26
1	368	361087	361800	gil216861	24K membrane protein [Pseudomonas aeruginosa]	52	32
1	376	368462	367695	gil147213	phnP protein [Escherichia coli]	52	47
1	381	373209	372412	gil467459	unknown [Bacillus subtilis]	52	28
1	437	418141	416768	gil1591425	hypothetical protein (GP:X91006_2) [Methanococcus jannaschii]	52	27
1	439	420801	420166	gil1674178	(AE000047) Mycoplasma pneumoniae, MG246 homolog, from M. genitalium [Mycoplasma pneumoniae]	52	40
1	474	443436	443798	gil1573287	aspartyl-tRNA synthetase (aspS) [Haemophilus influenzae]	52	35
1	583	555235	553802	gil496254	fibronectin/fibrinogen-binding protein [Streptococcus pyogenes]	52	29
1	759	715852	715610	gil397703	dihydroorotate dehydrogenase [Plasmodium falciparum]	52	30
1	797	751384	750674	gil1063419	S2 gene product [Borrelia burgdorferi]	52	27
1	820	776768	774852	gil580936	SpoVD [Bacillus subtilis]	52	33
1	869	820887	821516	gil1592298	ATP synthase, subunit D [Methanococcus jannaschii]	52	29
1	888	839581	838106	gil1151158	repeat organellar protein [Plasmodium falciparum]	52	30
1	916	862856	862110	gil1256625	putative [Bacillus subtilis]	52	25
1	67	83112	81610	gil587604	beta subunit RNA polymerase [Plasmodium falciparum]	51	29

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	150	147190	146360	gil520844	orf4 [Bacillus subtilis]	51	26
1	194	186516	185275	gil211931	3-hydroxy-3-methylglutaryl-CoA synthase [Gallus gallus]	51	29
1	300	288759	289676	gil142833	ORF2 [Bacillus subtilis]	51	33
1	371	362209	362874	gil1698880	protein antigen LmST11 [Leishmania major]	51	27
1	464	438943	439497	gil1591434	chromate resistance protein A [Methanococcus jannaschii]	51	29
1	819	772935	774842	gnlPIDle2390	AMP-binding protein [Brassica napus]	51	29
			57				
1	926	869257	869955	gil633996	a negative regulator of pho regulon [Pseudomonas aeruginosa]	51	26
1	45	54760	54062	gil505363	ORF2 [Salmonella typhimurium]	50	32
1	94	101155	102261	gil39995	phospho-N-acetylmuramoyl-pentapeptide- transferase [Bacillus subtilis]	50	29
1	118	118397	117096	gil1762996	RING-finger protein [Helicoverpa armigera nucleopolyhedrovirus]	50	25
1	155	151159	150506	gil893358	PgsA [Bacillus subtilis]	50	35
1	239	224187	224744	gil1303843	YqfV [Bacillus subtilis]	50	29
1	274	265044	264040	gnlPIDle2767	unknown [Mycobacterium tuberculosis]	50	32
			78				
1	287	276164	274710	gil147140	peptidase D [Escherichia coli]	50	28
1	310	299525	300778	gil1652202	ComE [Synechocystis sp.]	50	33
1	349	342477	341581	gnlPIDle2202	frameshift [Plasmodium falciparum]	50	30
			45				
1	457	435120	434509	gil144839	beta-galactosidase [Thermoanaerobacterium thermosulfurigenes]	50	29
1	479	448691	447948	gil580905	B.subtilis genes rpmH, mpA, 50kd, gidA and gidB [Bacillus subtilis]	50	32
1	680	640194	641039	gil882579	CG Site No. 29739 [Escherichia coli]	50	31
1	737	690152	690400	gil1086864	T03G11.2 gene product [Caenorhabditis elegans]	50	39
1	752	708130	709662	gil40162	murE gene product [Bacillus subtilis]	50	30
1	360	353288	354157	gil343314	involucrin [Saguinus oedipus]	49	20
1	44	54046	53216	gil505363	ORF2 [Salmonella typhimurium]	48	21
1	122	119896	120774	gil405908	yejE [Escherichia coli]	48	29
1	161	157504	156653	gil143213	putative [Bacillus subtilis]	48	23
1	316	305940	306995	gil1762962	FemA [Staphylococcus simulans]	48	28
1	459	436152	437315	gil1001478	hypothetical protein [Synechocystis sp.]	48	25

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	628	596267	596566	gill56218	putative [Caenorhabditis elegans]	48	32
1	694	655069	654452	gill574476	dedA protein (dedA) [Haemophilus influenzae]	48	22
1	731	686392	686129	gill915207	gastric mucin [Sus scrofa]	48	27
1	893	844951	843476	gnlPIDle220245	frameshift [Plasmodium falciparum]	48	32
1	62	74673	72196	gill766042	outer membrane protein [Neisseria gonorrhoeae]	47	30
1	103	107896	108780	gill256885	P24A protein (unknown function) (Swiss Prot. accession number P32802) [Saccharomyces cerevisiae]	47	27
1	187	181111	180215	gill184118	mevalonate kinase [Methanobacterium thermoautotrophicum]	47	30
1	204	195930	196616	gill573552	phosphoglycolate phosphatase, chromosomal (SP:P40852) [Haemophilus influenzae]	47	21
1	265	251835	251098	gill209847	repeat motif-containing gene [Borrelia burgdorferi]	47	30
1	334	325837	326529	gill591893	uridylylate kinase [Methanococcus jannaschii]	47	26
1	356	349581	349991	gill849173	Probable essential component of the nucleoskeleton (Swiss Prot. accession number P32380) [Saccharomyces cerevisiae]	47	27
1	490	460559	459834	gill592264	type I restriction enzyme [Methanococcus jannaschii]	47	34
1	526	499992	499264	gill710551	ankyrin 3 [Mus musculus]	47	29
1	577	549541	548390	gnlPIDle220245	frameshift [Plasmodium falciparum]	47	27
1	744	701189	699441	gnlPIDle160436	orfA gene product [Borrelia burgdorferi]	47	23
1	755	713050	710765	pirS41649IS41649	DNA polymerase - Plasmodium falciparum	47	22
1	761	717229	718392	gill500309	M. jannaschii predicted coding region MJ1428 [Methanococcus jannaschii]	47	37
1	813	767745	768737	pirG64100IG64100	membrane fusion protein (mtrC) homolog - Haemophilus influenzae (strain Rd KW20)	47	23
1	824	779587	780546	gill687844	contains TPR domain-like repeats [Caenorhabditis elegans]	47	28
1	881	834283	833015	gill574393	H. influenzae predicted coding region HI1548 [Haemophilus influenzae]	47	24
1	886	837236	836199	gill887563	serine/threonine-protein kinase [Plasmodium falciparum]	47	30
1	47	57001	55880	gill652686	hypothetical protein [Synechocystis sp.]	46	23
1	160	156659	156171	gill3261	ORF4 protein (AA 1-156) [Paramecium aurelia]	46	30

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	249	231765	232829	gil1142681	Lpp38 [Pasteurella haemolytica]	46	28
1	329	323695	322838	gil562039	NADH dehydrogenase, subunit 2 [Acanthamoeba castellanii]	46	20
1	336	329090	327303	gil457146	rhostry protein [Plasmodium yoelii]	46	18
1	442	422511	421747	gil1591598	hypothetical protein (GP:U19364_6) [Methanococcus jannaschii]	46	27
1	452	428632	429375	piriS41649IS41649	DNA polymerase - Plasmodium falciparum	46	21
1	573	545081	545596	gil1022328	Four tandem repeats of a DNA-binding domain known as the AT-hook are found at the carboxy terminus of Card. This protein has been purified and found to bind in vitro to a promoter region [Myxococcus xanthus]	46	28
1	617	586903	587865	gnlIPIDle33329	ND5 protein [Ascaris suum]	46	29
1	708	668290	666710	gil1573271	apolipoprotein N-acyltransferase (cute) [Haemophilus influenzae]	46	32
1	790	741189	740008	gil458015	TpN50 precursor [Treponema pallidum]	46	29
1	892	843474	841147	gil1574537	outer membrane integrity protein (tolA) [Haemophilus influenzae]	46	19
1	903	853463	852741	gil806562	nebulin [Homo sapiens]	46	26
1	968	908917	909948	gil438455	possible N-terminal signal sequence; mature protein may be membrane-anchored and start at Cys-17. 17.5% identity over 354-aa overlap with Candida pelliculosa beta-glucosidase.; putative [Bacillus subtilis]	46	29
1	208	197467	198516	gil220578	open reading frame [Mus musculus]	45	27
1	462	438197	438949	gil687689	similar to a chromate resistance protein (ChrA) from A. eutrophus, Swiss-Prot Accession Number P17551 [Synechococcus sp.]	45	23
1	742	698657	695295	gil1499043	M. jannaschii predicted coding region MJ0263 [Methanococcus jannaschii]	45	23
1	90	99196	98756	gil303895	ORF 8: This ORF is required for the secretion of IpaB, IpaC and IpaD [Plasmid pMYSH6000]	44	26
1	253	235698	234343	gil143245	Na+/H+ antiporter [Bacillus firmus]	44	26
1	709	668406	670430	gnlIPIDle236483	F54G8.4 [Caenorhabditis elegans]	44	31
1	850	802490	801045	gnlIPIDle220245	frameshift [Plasmodium falciparum]	44	25
1	458	436119	435118	gil1303799	YqeN [Bacillus subtilis]	43	19

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

1	859	810560	809967	pirS17998IS1 7998	gene COX1 intron 4 protein - yeast (<i>Kluyveromyces marxianus</i> var. <i>lactis</i>) mitochondrion (SGC2)	43	30
1	935	881179	879701	gil1045905	no score generated - score shown is bogus [<i>Mycoplasma genitalium</i>]	43	27
1	319	311250	309877	gil1591425	hypothetical protein (GP:X91006.2) [<i>Methanococcus jannaschii</i>]	42	19
1	618	587863	588672	gil1045801	hypothetical protein (SP:P32720) [<i>Mycoplasma genitalium</i>]	42	24
1	625	593472	594572	gil343962	VAR1 protein [<i>Candida glabrata</i>]	42	25
1	93	100191	101021	gil413976	ipa-52r gene product [<i>Bacillus subtilis</i>]	41	26
1	574	545523	546581	gnlPIDle16326	MURF2 protein (AA 1-348) [<i>Crithidia fasciculata</i>]	41	26
1	740	693458	692403	gil1151158	repeat organellar protein [<i>Plasmodium chabaudi</i>]	41	20
1	3	5792	6796	gil1256888	Similar to chromosome segregation protein Smc1p of <i>S. cerevisiae</i> (GenBank accession number L00602), chromosome segregation protein Cut3p of <i>S. pombe</i> (Swiss Prot. accession number P41004), and <i>C. elegans</i> hypothetical proteins R13G10.1 (GenBank)	40	24
1	228	214440	214742	gil1150836	neural specific DNA binding protein [<i>Xenopus laevis</i>]	40	28
1	318	309735	308377	gil1591425	hypothetical protein (GP:X91006.2) [<i>Methanococcus jannaschii</i>]	40	19
1	453	431037	429700	gil499647	[<i>Mus musculus</i> (strain C3H/RL) ORF mRNA, complete cds.], gene product [<i>Mus musculus</i>]	40	24
1	795	747813	749516	gil304179	wall-associated protein [<i>Bacillus subtilis</i>]	40	35
1	966	907336	905528	gil1151158	repeat organellar protein [<i>Plasmodium chabaudi</i>]	40	26

TABLE 2.
Borrelia burgdorferi - Coding regions containing known proteins

Contig ID	Orf ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	HSP nt length
1	69	86349	85018	gbL32144l	Borrelia burgdorferi peptidyl-tRNA hydrolase-like protein (pth) gene homologue, complete cds	100	220
1	70	86918	86340	gbL32144l	Borrelia burgdorferi peptidyl-tRNA hydrolase-like protein (pth) gene homologue, complete cds	100	579
1	71	87573	86911	gbL32144l	Borrelia burgdorferi peptidyl-tRNA hydrolase-like protein (pth) gene homologue, complete cds	100	129
1	124	123885	121759	gbM60802l	B.burgdorferi immunogen gene, 5' flank	99	2127
1	126	127421	125700	embIX91965l BBATPBP	B.burgdorferi abp gene	97	284
1	137	136332	139151	gbL31424l	Borrelia burgdorferi (clone BbK3.11) phoA fusion protein gene, partial cds	98	248
1	138	138676	138515	gbL31424l	Borrelia burgdorferi (clone BbK3.11) phoA fusion protein gene, partial cds	96	60
1	165	160705	159932	gbU17591l	Borrelia burgdorferi primary sigma factor (rpoD) gene, complete cds	100	774
1	166	162604	160703	gbU17591l	Borrelia burgdorferi primary sigma factor (rpoD) gene, complete cds	100	1902
1	167	162835	162602	gbU17591l	Borrelia burgdorferi primary sigma factor (rpoD) gene, complete cds	99	232
1	168	164397	162811	gbU17591l	Borrelia burgdorferi primary sigma factor (rpoD) gene, complete cds	99	1216
1	210	198495	199028	gbU61498l	Borrelia burgdorferi CheA (cheA) gene, partial cds, CheW (cheW), CheX (cheX) and CheY (cheY) genes, complete cds	98	127
1	211	199527	199069	gbU61498l	Borrelia burgdorferi CheA (cheA) gene, partial cds, CheW (cheW), CheX (cheX) and CheY (cheY) genes, complete cds	99	459
1	212	200067	199549	gbU61498l	Borrelia burgdorferi CheA (cheA) gene, partial cds, CheW (cheW), CheX (cheX) and CheY (cheY) genes, complete cds	99	519
1	213	201455	200046	gbU61498l	Borrelia burgdorferi CheA (cheA) gene, partial	99	1410

TABLE 2. *Borrelia burgdorferi* - Coding regions containing known proteins

							cds, CheW (cheW), CheX (cheX) and CheY (cheY) genes, complete cds		
1	214	204053	201453	gbU289621			<i>Borrelia burgdorferi</i> histidine kinase (cheA) gene, complete cds	99	2601
1	215	205173	204115	gbU629001			<i>Borrelia burgdorferi</i> flagellar filament outsheath protein (flaA) gene, complete cds, and chemotaxis histidine kinase (cheA) gene, partial cds	98	1059
1	216	205780	205220	gbU629001			<i>Borrelia burgdorferi</i> flagellar filament outsheath protein (flaA) gene, complete cds, and chemotaxis histidine kinase (cheA) gene, partial cds	98	277
1	235	220591	220232	embIX651391 BBHSP60			<i>B. burgdorferi</i> hsp60 gene for 60kDa heat shock protein	98	139
1	236	222288	220594	embIX651391 BBHSP60			<i>B. burgdorferi</i> hsp60 gene for 60kDa heat shock protein	99	1695
1	237	222396	224114	embIX540591 BBGROEL			<i>Borrelia burgdorferi</i> groEL gene for a common antigen	94	416
1	241	227310	225754	gbL314171			<i>Borrelia burgdorferi</i> (clone BbK1.4) phoA fusion protein gene, partial cds	99	231
1	295	281749	283629	embIX877251 BBDNA66K D			<i>B. burgdorferi</i> p66 gene for 66kDa protein	100	1881
1	296	284438	283683	gbIM584311			<i>Borrelia burgdorferi</i> PCR target sequence	93	356
1	315	303735	305942	gbL321461			<i>Borrelia burgdorferi</i> methionyl tRNA synthetase (metG) gene, partial cds	99	345
1	337	330246	329077	gbU602361			<i>Borrelia burgdorferi</i> response regulator gene, partial cds	100	191
1	338	332455	330299	gbL399651			<i>Borrelia burgdorferi</i> histidine kinase (cheA) gene, complete cds	99	1361
1	344	336894	335473	gbU518781			<i>Borrelia burgdorferi</i> phosphotransferase enzyme II (crr) gene, hsp90 (hptg) gene, complete cds	100	95
1	345	336851	338830	gbU518781			<i>Borrelia burgdorferi</i> phosphotransferase enzyme	99	1980

TABLE 2. *Borrelia burgdorferi* - Coding regions containing known proteins

						II (crr) gene, hsp90 (hptg) gene, complete cds		
1	346	339458	338868	gblU518781		<i>Borrelia burgdorferi</i> phosphotransferase enzyme II (crr) gene, hsp90 (hptg) gene, complete cds	100	591
1	388	378955	379590	gblM968471		<i>Borrelia burgdorferi</i> GrpE protein homologue gene, DnaK protein homologue gene, and DnaJ protein homologue gene, complete cds's	100	636
1	389	379566	381521	emblX676461 BBHSPRO		<i>B. burgdorferi</i> dnaK gene for heat-shock protein	100	1956
1	390	381512	381943	gblM979141		<i>Borrelia burgdorferi</i> DnaJ gene, complete cds	97	424
1	391	381907	382617	gblM968471		<i>Borrelia burgdorferi</i> GrpE protein homologue gene, DnaK protein homologue gene, and DnaJ protein homologue gene, complete cds's	100	687
1	392	382656	383360	gblM968471		<i>Borrelia burgdorferi</i> GrpE protein homologue gene, DnaK protein homologue gene, and DnaJ protein homologue gene, complete cds's	95	144
1	393	383005	382688	gblM968471		<i>Borrelia burgdorferi</i> GrpE protein homologue gene, DnaK protein homologue gene, and DnaJ protein homologue gene, complete cds's	95	144
1	394	384408	383416	gblU829781		<i>Borrelia burgdorferi</i> phenylalanyl-tRNA synthetase alpha subunit (pheS), phenylalanyl-tRNA synthetase beta subunit (pheT) and thioredoxin reductase (trxB) genes, complete cds	99	956
1	395	384799	384467	gblU829781		<i>Borrelia burgdorferi</i> phenylalanyl-tRNA synthetase alpha subunit (pheS), phenylalanyl-tRNA synthetase beta subunit (pheT) and thioredoxin reductase (trxB) genes, complete cds	99	292
1	396	386169	384733	gblU829781		<i>Borrelia burgdorferi</i> phenylalanyl-tRNA synthetase alpha subunit (pheS), phenylalanyl-tRNA synthetase beta subunit (pheT) and thioredoxin reductase (trxB) genes, complete cds	99	1416
1	397	387733	386144	gblU829781		<i>Borrelia burgdorferi</i> phenylalanyl-tRNA synthetase alpha subunit (pheS), phenylalanyl-tRNA synthetase beta subunit (pheT) and thioredoxin reductase (trxB) genes, complete cds	99	1220

TABLE 2.
Borrelia burgdorferi - Coding regions containing known proteins

1	398	394257	387727	gbU829781	Borrelia burgdorferi phenylalanyl-tRNA synthetase alpha subunit (pheS), phenylalanyl-tRNA synthetase beta subunit (pheT) and thioredoxin reductase (trxB) genes, complete cds	99	230
1	421	408559	407981	gbM286811	B. burgdorferi promoter region DNA	98	152
1	427	411019	410132	gbU781931	Borrelia burgdorferi tuf-s10 operon: elongation factor (tuf), ribosomal proteins S10 (rpsJ), L3 (rplC), L4 (rplD), L23 (rplW), L2 (rplB), S19 (rpsS), and L22 (rplV) genes, complete cds, and S3 (rpsC) gene, partial cds	99	287
1	428	411388	411017	gbU781931	Borrelia burgdorferi tuf-s10 operon: elongation factor (tuf), ribosomal proteins S10 (rpsJ), L3 (rplC), L4 (rplD), L23 (rplW), L2 (rplB), S19 (rpsS), and L22 (rplV) genes, complete cds, and S3 (rpsC) gene, partial cds	96	357
1	429	411676	411386	gbU781931	Borrelia burgdorferi tuf-s10 operon: elongation factor (tuf), ribosomal proteins S10 (rpsJ), L3 (rplC), L4 (rplD), L23 (rplW), L2 (rplB), S19 (rpsS), and L22 (rplV) genes, complete cds, and S3 (rpsC) gene, partial cds	99	291
1	430	412531	411674	gbU781931	Borrelia burgdorferi tuf-s10 operon: elongation factor (tuf), ribosomal proteins S10 (rpsJ), L3 (rplC), L4 (rplD), L23 (rplW), L2 (rplB), S19 (rpsS), and L22 (rplV) genes, complete cds, and S3 (rpsC) gene, partial cds	98	858
1	431	412852	412529	gbU781931	Borrelia burgdorferi tuf-s10 operon: elongation factor (tuf), ribosomal proteins S10 (rpsJ), L3 (rplC), L4 (rplD), L23 (rplW), L2 (rplB), S19 (rpsS), and L22 (rplV) genes, complete cds, and S3 (rpsC) gene, partial cds	99	324
1	432	413487	412846	gbU781931	Borrelia burgdorferi tuf-s10 operon: elongation factor (tuf), ribosomal proteins S10 (rpsJ), L3 (rplC), L4 (rplD), L23 (rplW), L2 (rplB), S19 (rpsS), and L22 (rplV) genes, complete cds, and S3 (rpsC) gene, partial cds	99	642

TABLE 2. *Borrelia burgdorferi* - Coding regions containing known proteins

1	433	414117	413485	gbIU781931	S3 (rpsC) gene, partial cds		99	633
1	434	414464	414141	gbIU781931	<i>Borrelia burgdorferi</i> tuf-s10 operon: elongation factor (tuf), ribosomal proteins S10 (rpsJ), L3 (rplC), L4 (rplD), L23 (rplW), L2 (rplB), S19 (rpsS), and L22 (rplV) genes, complete cds, and S3 (rpsC) gene, partial cds		100	324
1	435	415714	414503	gbL231251	<i>Borrelia burgdorferi</i> tuf-s10 operon: elongation factor (tuf), ribosomal proteins S10 (rpsJ), L3 (rplC), L4 (rplD), L23 (rplW), L2 (rplB), S19 (rpsS), and L22 (rplV) genes, complete cds, and S3 (rpsC) gene, partial cds		100	1212
1	481	450681	450310	gbIU045271	<i>Borrelia burgdorferi</i> elongation factor EF-Tu (tuf) gene, complete cds		100	148
1	482	450820	450650	gbIU045271	<i>Borrelia burgdorferi</i> 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds		100	171
1	483	451208	450897	gbIU045271	<i>Borrelia burgdorferi</i> 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds		100	312
1	484	451288	451467	gbIU045271	<i>Borrelia burgdorferi</i> 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds		100	180

Borrelia burgdorferi - Coding regions containing known proteins

TABLE 2.

							DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds		
1	485	452456	451287	gblU045271			Borrelia burgdorferi 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds	99	1170
1	486	454181	452685	gblU045271			Borrelia burgdorferi 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds	100	1497
1	487	454315	456237	gblU045271			Borrelia burgdorferi 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds	98	904
1	488	456228	458681	embIZ12165IB BGYRAG			B. burgdorferi gyrA gene encoding DNA gyrase subunit A (partial)	96	289
1	496	463825	464394	gblU03396l			Borrelia burgdorferi B31 Ala-tRNA (alaT), Ile-tRNA (ileT), 16S rRNA, 23S rRNA (rrlA and rrlB), and 5S rRNA (rrfA and rrfB) genes, complete sequence	100	570
1	497	466650	466958	gblU03396l			Borrelia burgdorferi B31 Ala-tRNA (alaT), Ile-tRNA (ileT), 16S rRNA, 23S rRNA (rrlA and rrlB), and 5S rRNA (rrfA and rrfB) genes, complete sequence	96	210
1	498	467437	468033	gblU03396l			Borrelia burgdorferi B31 Ala-tRNA (alaT), Ile-tRNA (ileT), 16S rRNA, 23S rRNA (rrlA and rrlB), and 5S rRNA (rrfA and rrfB) genes, complete sequence	99	209

TABLE 2. *Borrelia burgdorferi* - Coding regions containing known proteins

1	499	468167	468433	gbIU033961	complete sequence	98	267
					<i>Borrelia burgdorferi</i> B31 Ala-tRNA (alaT), Ile-tRNA (ileT), 16S rRNA, 23S rRNA (rrlA and rrlB), and 5S rRNA (rrfA and rrfB) genes, complete sequence		
1	500	468391	468999	gbIU033961	complete sequence	95	386
					<i>Borrelia burgdorferi</i> B31 Ala-tRNA (alaT), Ile-tRNA (ileT), 16S rRNA, 23S rRNA (rrlA and rrlB), and 5S rRNA (rrfA and rrfB) genes, complete sequence		
1	501	470714	470445	gbIM883301	<i>Borrelia burgdorferi</i> 23S ribosomal RNA gene	100	270
1	502	475597	480090	gbIU033961	<i>Borrelia burgdorferi</i> B31 Ala-tRNA (alaT), Ile-tRNA (ileT), 16S rRNA, 23S rRNA (rrlA and rrlB), and 5S rRNA (rrfA and rrfB) genes, complete sequence	97	131
1	535	505532	509017	gbIL484881	<i>Borrelia burgdorferi</i> RNA polymerase beta subunit (rpoB) gene, complete cds, RNA polymerase beta' subunit (rpoC) gene, 5' end of cds	98	2490
1	536	509015	513166	gbIL484881	<i>Borrelia burgdorferi</i> RNA polymerase beta subunit (rpoB) gene, complete cds, RNA polymerase beta' subunit (rpoC) gene, 5' end of cds	97	76
1	538	513606	514106	gbIU354501	<i>Borrelia burgdorferi</i> membrane protein D (bmpD) gene, complete cds	100	82
1	539	514120	515229	gbIU354501	<i>Borrelia burgdorferi</i> membrane protein D (bmpD) gene, complete cds	99	1110
1	540	515472	516605	gbIU499381	<i>Borrelia burgdorferi</i> potential virulence gene cluster membrane proteins BmpC (bmpC) and BmpA (bmpA), BmpB protein (bmpB), putative protein 4, Mg ion transporter MgtE (mgtE), protein kinase C1 inhibitor PKCI (pkci) genes, complete cds	99	1134
1	541	516641	517666	gbIL241941	<i>Borrelia burgdorferi</i> immunodominant antigen P39 gene, complete cds	99	1026

TABLE 2. *Borrelia burgdorferi* - Coding regions containing known proteins

1	542	517732	518256	gblL350501	<i>Borrelia burgdorferi</i> (clone pB46) membrane lipoprotein A (bnpA) gene, 3' end, membrane lipoprotein (bnpB) gene, 5' end	98	457
1	543	518168	518779	gblL241941	<i>Borrelia burgdorferi</i> immunodominant antigen P39 gene, complete cds	99	606
1	544	518856	520316	gblU499381	<i>Borrelia burgdorferi</i> potential virulence gene cluster membrane proteins BmpC (bnpC) and BmpA (bnpA), BmpB protein (bnpB), putative protein 4, Mg ion transporter MgtE (mgtE), protein kinase C1 inhibitor PKCI (pkci) genes, complete cds	99	1461
1	545	520349	521734	gblU499381	<i>Borrelia burgdorferi</i> potential virulence gene cluster membrane proteins BmpC (bnpC) and BmpA (bnpA), BmpB protein (bnpB), putative protein 4, Mg ion transporter MgtE (mgtE), protein kinase C1 inhibitor PKCI (pkci) genes, complete cds	99	1386
1	546	521752	522204	gblU499381	<i>Borrelia burgdorferi</i> potential virulence gene cluster membrane proteins BmpC (bnpC) and BmpA (bnpA), BmpB protein (bnpB), putative protein 4, Mg ion transporter MgtE (mgtE), protein kinase C1 inhibitor PKCI (pkci) genes, complete cds	100	453
1	547	522168	522893	gblU499381	<i>Borrelia burgdorferi</i> potential virulence gene cluster membrane proteins BmpC (bnpC) and BmpA (bnpA), BmpB protein (bnpB), putative protein 4, Mg ion transporter MgtE (mgtE), protein kinase C1 inhibitor PKCI (pkci) genes, complete cds	98	130
1	559	535086	534772	embIX787081 BBYSC1	<i>B. burgdorferi</i> (ZS7) YSC1-like gene	99	314
1	560	536461	535058	embIX787081 BBYSC1	<i>B. burgdorferi</i> (ZS7) YSC1-like gene	100	1404
1	561	536545	537144	embIX708261	<i>B. burgdorferi</i> gene for lipoprotein	100	600

TABLE 2. Borrelia burgdorferi - Coding regions containing known proteins

1	562	537652	537191	BBLA7 embIX708261 BBLA7	B. burgdorferi gene for lipoprotein	100	57
1	563	539695	537665	gblM900841	Borrelia burgdorferi 22 kD antigen	100	786
1	564	537705	537968	gblM900841	Borrelia burgdorferi 22 kD antigen	100	264
1	565	538395	538757	gblM900841	Borrelia burgdorferi 22 kD antigen	100	56
1	606	574092	572497	gblU291431	Borrelia burgdorferi periplasmic substrate-binding protein homolog (p30) gene, complete cds	92	805
1	607	575817	574204	gblU291431	Borrelia burgdorferi periplasmic substrate-binding protein homolog (p30) gene, complete cds	100	84
1	616	585458	586936	gblL314221	Borrelia burgdorferi (clone Bb2.13) phoA fusion protein gene, partial cds	100	354
1	629	596586	597983	gblU437391	Borrelia burgdorferi fmsmid clone 31, complete sequence	100	1185
1	630	597967	599052	embIX966851 BBCDG	B. burgdorferi cell division genes	99	912
1	631	599050	600153	embIX966851 BBCDG	B. burgdorferi cell division genes	99	1104
1	632	600183	600932	embIX964331 BBFTSWQA	B. burgdorferi ftsW, ftsQ & ftsA genes	99	750
1	633	600905	602173	gblU437391	Borrelia burgdorferi fmsmid clone 31, complete sequence	100	1269
1	634	602171	603394	gblU437391	Borrelia burgdorferi fmsmid clone 31, complete sequence	100	1224
1	635	603392	604087	gblU437391	Borrelia burgdorferi fmsmid clone 31, complete sequence	100	696
1	636	604085	605041	gblL763031	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flfEFGHI, flbABC genes, complete cds	98	712
1	637	605039	605599	gblU437391	Borrelia burgdorferi fmsmid clone 31, complete sequence	100	561

TABLE 2.
Borrelia burgdorferi - Coding regions containing known proteins

1	638	605535	606938	embIX966851 BBCDG	B. burgdorferi cell division genes	97	1404
1	639	606936	607379	gbIU43739I	Borrelia burgdorferi fesmid clone 31, complete sequence	100	444
1	640	607382	607861	gbIU43739I	Borrelia burgdorferi fesmid clone 31, complete sequence	100	480
1	641	607831	608208	gbIL76303I	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flieFGHI, flbABC genes, complete cds	100	378
1	642	608163	609932	gbIL76303I	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flieFGHI, flbABC genes, complete cds	100	1770
1	643	609930	610982	gbIU43739I	Borrelia burgdorferi fesmid clone 31, complete sequence	100	1053
1	644	610961	611917	gbIU43739I	Borrelia burgdorferi fesmid clone 31, complete sequence	100	957
1	645	611915	613246	gbIL76303I	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flieFGHI, flbABC genes, complete cds	99	1332
1	646	613222	613674	gbIL76303I	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flieFGHI, flbABC genes, complete cds	99	453
1	647	613655	614284	gbIU43739I	Borrelia burgdorferi fesmid clone 31, complete sequence	100	630
1	648	614250	615470	gbIL76303I	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flieFGHI, flbABC genes, complete cds	99	1221
1	649	615481	615927	gbIL76303I	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flieFGHI, flbABC genes, complete cds	100	447
1	650	615911	617260	gbIU43739I	Borrelia burgdorferi fesmid clone 31, complete sequence	100	1350
1	651	617277	617507	gbIU43739I	Borrelia burgdorferi fesmid clone 31, complete sequence	100	231

TABLE 2.
Borrelia burgdorferi - Coding regions containing known proteins

1	652	617498	618286	gblU43739	Borrelia burgdorferi fesmid clone 31, complete sequence	100	789
1	653	618280	619068	gblU43739	Borrelia burgdorferi fesmid clone 31, complete sequence	100	789
1	654	619066	619653	gblU43739	Borrelia burgdorferi fesmid clone 31, complete sequence	100	588
1	655	619688	620749	gblU43739	Borrelia burgdorferi fesmid clone 31, complete sequence	100	1062
1	656	620789	621136	gblU43739	Borrelia burgdorferi fesmid clone 31, complete sequence	100	348
1	657	621114	621755	gblU43739	Borrelia burgdorferi fesmid clone 31, complete sequence	100	642
1	658	621742	622530	gblL75945	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flilL, flilM, fliz, flagellar export apparatus (fliPQR, flhB), flhF, flbE genes	99	789
1	659	622028	621822	gblL75945	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flilL, flilM, fliz, flagellar export apparatus (fliPQR, flhB), flhF, flbE genes	100	207
1	660	622515	622802	gblL75945	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flilL, flilM, fliz, flagellar export apparatus (fliPQR, flhB), flhF, flbE genes	100	288
1	661	622811	623623	gblL75945	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flilL, flilM, fliz, flagellar export apparatus (fliPQR, flhB), flhF, flbE genes	99	813
1	662	623007	622819	gblL75945	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flilL, flilM, fliz, flagellar export apparatus (fliPQR, flhB), flhF, flbE genes	100	189
1	663	623706	623458	gblL75945	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flilL,	100	249

TABLE 2. *Borrelia burgdorferi* - Coding regions containing known proteins

							fliM, fliZ, flagellar export apparatus (flpQR, flhB), flhF, flbE genes			
1	664	623608	624741	gblL759451			Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), fliL, fliM, fliZ, flagellar export apparatus (flpQR, flhB), flhF, flbE genes	99		1134
1	665	624735	626843	gblL759451			Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), fliL, fliM, fliZ, flagellar export apparatus (flpQR, flhB), flhF, flbE genes	100		2109
1	666	626841	628013	gblU437391			Borrelia burgdorferi ftsmid clone 31, complete sequence	100		1173
1	667	627998	628912	gblU437391			Borrelia burgdorferi ftsmid clone 31, complete sequence	99		816
1	668	629151	628807	gblU437391			Borrelia burgdorferi ftsmid clone 31, complete sequence	100		345
1	669	628910	629398	gblU437391			Borrelia burgdorferi ftsmid clone 31, complete sequence	100		489
1	670	629371	631305	gblU437391			Borrelia burgdorferi ftsmid clone 31, complete sequence	100		1935
1	671	631314	631634	gblU437391			Borrelia burgdorferi ftsmid clone 31, complete sequence	100		286
1	676	636891	635476	gblM286821			B. burgdorferi promoter element DNA	100		78
1	687	646982	649420	gblL772161			Borrelia burgdorferi (strain B31) protease (lon) gene, complete cds	99		2439
1	688	651760	649409	gblL772161			Borrelia burgdorferi (strain B31) protease (lon) gene, complete cds	100		274
1	711	671567	672412	gblU356731			Borrelia burgdorferi OrfR gene, partial cds, and S20, Hbb, OrfH and Rho genes, complete cds	99		542
1	712	672418	672744	gblU356731			Borrelia burgdorferi OrfR gene, partial cds, and S20, Hbb, OrfH and Rho genes, complete cds	100		327
1	713	672751	673083	gblU486511			Borrelia burgdorferi PIG histone-like protein Hbbu (hbb) gene, complete cds	100		327

TABLE 2. *Borrelia burgdorferi* - Coding regions containing known proteins

1	714	673081	673491	gbU356731	<i>Borrelia burgdorferi</i> OrfR gene, partial cds, and S20, Hbb, OrfH and Rho genes, complete cds	99	411
1	715	673553	675118	gbU356731	<i>Borrelia burgdorferi</i> OrfR gene, partial cds, and S20, Hbb, OrfH and Rho genes, complete cds	99	1566
1	716	675164	675424	gbU356731	<i>Borrelia burgdorferi</i> OrfR gene, partial cds, and S20, Hbb, OrfH and Rho genes, complete cds	100	106
1	773	724171	723770	gbU629011	<i>Borrelia burgdorferi</i> thdF gene, partial cds, putative motility protein (flbF), flagellar hook associated proteins FlgK (flgK) and FlgL (flgL) genes, complete cds	97	67
1	774	723891	724181	gbU629011	<i>Borrelia burgdorferi</i> thdF gene, partial cds, putative motility protein (flbF), flagellar hook associated proteins FlgK (flgK) and FlgL (flgL) genes, complete cds	97	77
1	775	725456	724164	gbU629011	<i>Borrelia burgdorferi</i> thdF gene, partial cds, putative motility protein (flbF), flagellar hook associated proteins FlgK (flgK) and FlgL (flgL) genes, complete cds	99	780
1	776	727348	725441	gbU629011	<i>Borrelia burgdorferi</i> thdF gene, partial cds, putative motility protein (flbF), flagellar hook associated proteins FlgK (flgK) and FlgL (flgL) genes, complete cds	99	1841
1	777	727854	727336	gbU629011	<i>Borrelia burgdorferi</i> thdF gene, partial cds, putative motility protein (flbF), flagellar hook associated proteins FlgK (flgK) and FlgL (flgL) genes, complete cds	99	519
1	778	727908	729308	embIX956691 BBTHDFGID	<i>B. burgdorferi</i> thdF and gidA genes	98	1185
1	779	729284	731176	embZ121601B BGIDAG	<i>B. burgdorferi</i> thdF, gidA and gidB genes	99	1893
1	780	731149	731799	embIX956681 BBGIDMOX R	<i>B. burgdorferi</i> gidA, gidB and moxR genes	98	381
1	781	731772	732848	embIX964341	<i>B. burgdorferi</i> gidB moxR genes and ORF	99	789

TABLE 2. Borrelia burgdorferi - Coding regions containing known proteins

					BBGIDBMO X			
1	782	732815	733738	embIX96434 BBGIDBMO X	B.burgdorferi gidB moxR genes and ORF	100	84	
1	798	752154	751372	gblU51878	Borrelia burgdorferi phosphotransferase enzyme II (crr) gene, hsp90 (hptg) gene, complete cds	100	57	
1	800	754266	753118	gblAF003354	Borrelia burgdorferi SecA (secA) gene, complete cds	97	67	
1	801	753992	754243	gblAF003354	Borrelia burgdorferi SecA (secA) gene, complete cds	96	50	
1	802	754283	757015	gblAF003354	Borrelia burgdorferi SecA (secA) gene, complete cds	99	2041	
1	803	756991	757641	gblAF003354	Borrelia burgdorferi SecA (secA) gene, complete cds	100	158	
1	806	759909	761930	gblU66699	Borrelia burgdorferi flagellar filament cap (filD) gene, complete cds and flagellin protein (flaB) gene, partial cds	98	1149	
1	807	762051	763067	embIX16833 BBFAA	Borrelia burgdorferi gene for flagellum-associated 41kD antigen (flagellin)	99	1017	
1	808	763194	764339	embIX63898 BBHYPP	B.burgdorferi DNA for hypothetical protein	99	1146	
1	809	764337	765245	embIX63898 BBHYPP	B.burgdorferi DNA for hypothetical protein	92	253	
1	826	783276	784400	gblU23457	Borrelia burgdorferi RecA (recA) gene, complete cds	99	1122	
1	827	784412	785182	gblU23457	Borrelia burgdorferi RecA (recA) gene, complete cds	82	476	
1	828	785142	785918	gblU23457	Borrelia burgdorferi RecA (recA) gene, complete cds	99	139	
1	907	855179	857182	gblU28760	Borrelia burgdorferi glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK), triosephosphate isomerase (TPI)	94	75	

TABLE 2. Borrelia burgdorferi - Coding regions containing known proteins

1	908	857228	858262	gbU28760	genes, complete cds		99	1035
1	909	858270	859463	gbU28760	Borrelia burgdorferi glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK), triosephosphate isomerase (TPI) genes, complete cds		99	1194
1	910	859315	860226	gbU28760	Borrelia burgdorferi glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK), triosephosphate isomerase (TPI) genes, complete cds		99	912
1	911	860224	860604	gbU57683	Borrelia burgdorferi sequence 3' to the triosephosphate isomerase (TPI) gene		97	183
1	912	860645	860316	gbU57683	Borrelia burgdorferi sequence 3' to the triosephosphate isomerase (TPI) gene		95	94
1	913	861447	860704	gbU57684	Borrelia burgdorferi uracil DNA glycosylase (UDG) gene, partial cds		92	294
1	914	861020	861397	gbU57684	Borrelia burgdorferi uracil DNA glycosylase (UDG) gene, partial cds		93	244
1	915	861439	862113	gbU57684	Borrelia burgdorferi uracil DNA glycosylase (UDG) gene, partial cds		96	128
1	930	874089	874859	gbL32861	Borrelia burgdorferi 1-acyl-sn-glycerol-3-phosphate acetyltransferase (plsC) gene, 3' end; topoisomerase IV beta-subunit (parE) gene, 5' end		99	408
1	931	874877	876679	gbL32861	Borrelia burgdorferi 1-acyl-sn-glycerol-3-phosphate acetyltransferase (plsC) gene, 3' end; topoisomerase IV beta-subunit (parE) gene, 5' end		100	252
1	943	887900	886758	embIY088851 BBRUVABH L	B. burgdorferi ruvA, ruvB and queA genes		98	293

TABLE 2. *Borrelia burgdorferi* - Coding regions containing known proteins

1	944	887965	888570	emb Y08885 BBRUVABH L	B. burgdorferi ruvA, ruvB and queA genes	99	606
1	945	888603	889658	emb Y08885 BBRUVABH L	B. burgdorferi ruvA, ruvB and queA genes	99	1056
1	946	889615	890271	emb Y08885 BBRUVABH L	B. burgdorferi ruvA, ruvB and queA genes	97	342
1	948	890719	892404	emb Y09140 BBPFPB	B. burgdorferi pfpB gene	99	1320
1	950	892893	893909	emb Y09142 BBYFII	B. burgdorferi yfiI gene	97	919
1	952	894973	895371	emb X97449 BBPRIAUDK	B. burgdorferi priA and udk genes	88	324
1	953	895308	895991	emb X97449 BBPRIAUDK	B. burgdorferi priA and udk genes	99	684
1	954	897976	895988	emb X97449 BBPRIAUDK	B. burgdorferi priA and udk genes	99	1989
1	955	898577	897963	emb X97449 BBPRIAUDK	B. burgdorferi priA and udk genes	100	152
1	956	899298	898555	emb Y09141 BBTRUA	B. burgdorferi truA gene	99	741

TABLE 3.
Borrelia burgdorferi - Putative coding regions of novel proteins not similar to know proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
1	1	2330	1134
1	2	3317	2934
1	8	11375	13021
1	9	11673	11386
1	10	12925	13629
1	11	13538	14146
1	17	25212	24700
1	18	25782	25357
1	19	26115	25870
1	21	27308	27051
1	22	29628	30458
1	29	40696	41217
1	30	41201	41992
1	31	42542	41985
1	32	42593	42982
1	34	44234	44031
1	38	48041	47079
1	41	49318	49617
1	43	53234	51810
1	50	59737	58208
1	58	68227	67733
1	65	79757	80404
1	66	81516	80401
1	75	89552	88353
1	82	93338	92766
1	85	95207	95854
1	104	108788	108621
1	105	109764	108943
1	108	112003	111599
1	113	114317	115846
1	114	114522	114316
1	119	118439	118927
1	121	119802	119599
1	125	125688	123967
1	129	128594	129235
1	135	136116	135259
1	136	136558	136298
1	139	139149	139559
1	141	140573	140121
1	143	141738	141412
1	145	142218	142060
1	146	142686	142342
1	154	150528	149074
1	158	153832	153981
1	163	158277	158474

Borrelia burgdorferi - Putative coding regions of novel proteins not similar to know proteins

1	171	168052	166205
1	176	171592	171038
1	186	179607	180089
1	189	182345	182046
1	191	182567	182773
1	199	192561	192716
1	205	196592	197476
1	218	207717	206752
1	219	207733	208437
1	221	209337	208915
1	222	209712	209335
1	231	217179	216025
1	238	223660	223418
1	240	224720	225724
1	242	227006	227275
1	248	231761	231501
1	251	232973	233308
1	252	233669	234004
1	254	235115	235456
1	258	241824	242198
1	261	248009	247773
1	269	256846	255872
1	276	265430	265158
1	279	266582	266298
1	281	268474	268280
1	286	274157	274384
1	292	280495	280274
1	294	281344	281042
1	298	287276	285714
1	303	292943	292644
1	304	293273	293037
1	305	294965	294648
1	308	299427	298699
1	309	299051	299212
1	326	320375	319785
1	327	320425	321036
1	331	324198	324413
1	339	332785	332459
1	341	333503	334138
1	342	334116	334739
1	343	334880	335446
1	350	342916	342443
1	351	344789	342897
1	363	357596	356931
1	367	361065	360859
1	370	362519	362196

Borrelia burgdorferi - Putative coding regions of novel proteins not similar to know proteins

1	374	366905	366114
1	377	368632	369537
1	378	369928	370560
1	379	370532	371353
1	382	375028	373193
1	383	375102	375542
1	387	378677	378198
1	400	394952	394722
1	401	396247	394937
1	403	397569	398327
1	406	399103	399294
1	436	416160	416570
1	445	424660	423950
1	446	425181	424642
1	450	428559	428200
1	451	428933	428619
1	455	432590	431628
1	461	437823	438092
1	463	438690	438313
1	466	440749	440222
1	470	441568	441350
1	471	442039	441614
1	472	442216	442037
1	473	442666	442262
1	476	445202	445017
1	493	462106	462519
1	494	462893	462549
1	504	482111	481035
1	505	481552	481800
1	509	483249	483668
1	512	484864	485157
1	516	489171	488527
1	519	492989	492375
1	520	493626	492997
1	521	494169	494864
1	524	497185	497385
1	525	497674	499254
1	527	500251	501294
1	528	501281	502156
1	558	533912	533667
1	568	541267	541491
1	571	544436	544257
1	572	544565	545068
1	578	549603	551198
1	580	551508	551657
1	581	552337	551513

Borrelia burgdorferi - Putative coding regions of novel proteins not similar to know proteins

1	585	556051	557271
1	590	561342	561139
1	591	561825	561520
1	592	562536	563360
1	596	565758	566519
1	599	568389	568682
1	602	568680	568856
1	605	570829	571167
1	609	576170	577093
1	612	581549	581091
1	614	582910	584013
1	619	589384	588674
1	624	592665	593465
1	626	594542	595405
1	672	631642	632175
1	677	636650	636892
1	678	637059	638078
1	681	640861	640412
1	686	644887	645207
1	689	649716	649961
1	690	650436	650735
1	691	650733	651056
1	693	653303	653689
1	705	664733	664918
1	707	665979	666770
1	718	679155	678391
1	721	680664	681047
1	722	681523	681849
1	724	681809	682171
1	727	682853	683272
1	734	687648	688067
1	739	691613	692290
1	751	707290	707718
1	763	719197	718904
1	764	720030	719257
1	769	722198	722482
1	783	733736	734647
1	785	735554	736618
1	787	737124	739184
1	792	742924	744801
1	799	753128	752655
1	811	766129	765980
1	812	766438	767772
1	815	770062	769790
1	818	771890	772282
1	831	788219	788836

Borrelia burgdorferi - Putative coding regions of novel proteins not similar to know proteins

1	832	788824	789615
1	838	793566	793414
1	840	794295	794119
1	844	796774	796586
1	852	803096	802908
1	858	809371	809970
1	864	816108	816497
1	865	816672	817283
1	866	817281	817838
1	872	823841	824836
1	876	828191	828739
1	877	828749	829147
1	879	831328	831714
1	880	831698	833005
1	885	836201	835677
1	890	841171	840590
1	891	840594	840860
1	899	849453	850148
1	902	851608	852687
1	918	862867	863109
1	920	864292	864705
1	923	865660	865346
1	925	868212	869273
1	928	871012	872580
1	933	878576	879166
1	939	884338	883268
1	940	884999	884325
1	949	892388	892924
1	957	900141	899296
1	958	900534	900139
1	959	901526	900510
1	962	902383	903258

TABLE 4.
Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident
69	1	291	4	gil146582	beta-lactamase [Escherichia coli]	100	98
69	2	692	240	gil344797	galactosidase fusion protein [unidentified]	100	99
26	3	1575	2093	gil458219	ORF 4 [Borrelia burgdorferi]	94	76
2	48	41836	41459	gil47453	ribosomal protein S12 [Streptococcus pneumoniae]	92	85
6	20	14234	12951	bbs1161785	60 kda antigen [Borrelia coriaceae, C053, ATCC 4338, Peptide, 514 aa] [Borrelia coriaceae]	88	67
52	5	1080	1652	gnlPID1e201250	ORF-D gene product [Borrelia burgdorferi]	88	74
152	1	337	26	gnlPID1e158979	orfA gene product [Borrelia burgdorferi]	86	75
71	2	1421	1128	gnlPID1e160437	orfD gene product [Borrelia burgdorferi]	85	46
131	1	381	674	gil458220	ORF 5 [Borrelia burgdorferi]	85	76
3	113	98152	97367	gil1591672	phosphate transport system ATP-binding protein [Methanococcus jannaschii]	84	65
2	107	108403	109485	gil882454	fructose 1,6-bisphosphate aldolase [Escherichia coli]	81	61
19	4	4059	4754	pirA34520A34520	29K calcium-binding protein, brain-specific - guinea pig (fragments)	81	56
20	9	6084	5791	gnlPID1e201249	ORF-C gene product [Borrelia burgdorferi]	81	72
2	52	49986	49600	pirA02771IR7MCML	ribosomal protein L7/L12 - Micrococcus luteus	80	67
14	1	3071	3	gil1522636	M. jannaschii predicted coding region MJEC502 [Methanococcus jannaschii]	80	60
29	2	218	409	gil1752736	gene required for phosphorylation of oligosaccharides/ has high homology with YJR061w [Saccharomyces cerevisiae]	80	37
32	2	719	925	gil433720	CDC25 [Homo sapiens]	80	73
100	1	2	946	gil1522636	M. jannaschii predicted coding region MJEC502 [Methanococcus jannaschii]	80	60

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

2	106	107148	108239	gnlIPIDle2881 24	glucose epimerase [Bacillus thuringiensis]	79	66
8	4	4878	4735	gil1543076	outer membrane porin protein Oms28 precursor [Borrelia burgdorferi]	79	69
2	55	51661	51218	gil587583	ribosomal protein L11 [Thermus aquaticus thermophilus]	78	58
4	54	39290	38742	gnlIPIDle1604 37	orfD gene product [Borrelia burgdorferi]	78	57
5	46	27416	27177	gnlIPIDle2532 11	ORF YDL065c [Saccharomyces cerevisiae]	78	57
7	4	2382	2966	gnlIPIDle2012 48	ORF-B gene product [Borrelia burgdorferi]	78	60
19	5	5107	4943	gil882579	CG Site No. 29739 [Escherichia coli]	78	42
78	1	1	171	gnlIPIDle2012 49	ORF-C gene product [Borrelia burgdorferi]	78	60
105	2	503	742	gnlIPIDle2532 11	ORF YDL065c [Saccharomyces cerevisiae]	78	57
2	30	24917	23697	gil143795	transfer RNA-Tyr synthetase [Bacillus subtilis]	77	52
6	34	22722	24080	gil466474	cellobiose phosphotransferase enzyme II" [Bacillus stearothermophilus]	77	50
8	1	688	536	gil1017809	similar to dihydropyridine-sensitive I-type, skeletal muscle calcium channel alpha-1 subunit (SP:CIC1_RABIT, P07293) [Caenorhabditis elegans]	77	55
3	91	81071	82183	gil467376	unknown [Bacillus subtilis]	76	58
11	1	208	2	gil1065989	(pos:5995..5997,aa:Met) [Bacillus subtilis]	76	56
68	1	605	3	gnlIPIDle1589 80	orfC gene product [Borrelia burgdorferi]	76	56
2	9	8488	6674	pirIC30010IC3 0010	hypothetical ORF-6 protein - Sauroleishmania tarentolae mitochondrion (SGC6)	75	50
2	37	31639	32163	gil1573470	H. influenzae predicted coding region HI0491 [Haemophilus influenzae]	75	55
2	56	52261	51701	gil396321	nusG [Escherichia coli]	75	56

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

20	1	653	414	gil520778	protein p23 [Borrelia burgdorferi]	75	62
20	3	2437	1652	gnlIPIDle2012 49	ORF-C gene product [Borrelia burgdorferi]	75	60
58	1	856	62	gnlIPIDle2012 49	ORF-C gene product [Borrelia burgdorferi]	75	37
68	3	1153	578	gil458217	ORF 2 [Borrelia burgdorferi]	75	55
117	1	744	388	gil520783	unknown [Borrelia burgdorferi]	75	42
130	1	1	684	gnlIPIDle1604 36	orfA gene product [Borrelia burgdorferi]	75	58
2	36	30506	31693	gil1020317	S-adenosylmethionine synthetase [Staphylococcus aureus]	74	57
2	109	111301	109871	gil396501	aspartyl-tRNA synthetase [Thermus aquaticus thermophilus]	74	52
3	101	92143	91103	gil1651962	hypothetical protein [Synechocystis sp.]	74	49
20	5	4080	2974	gnlIPIDle1589 79	orfA gene product [Borrelia burgdorferi]	74	56
36	2	468	1253	gnlIPIDle1589 84	orfC gene product [Borrelia burgdorferi]	74	59
42	1	396	719	gil1655798	CdsK [Borrelia burgdorferi]	74	58
2	10	6810	7022	gil406135	glycoprotein 120 [Simian immunodeficiency virus]	73	53
2	29	23695	21395	gil511145	hemolysin [Serpulina hyodysenteriae]	73	52
3	56	44789	44262	gil1595810	type-I signal peptidase SpsB [Staphylococcus aureus]	73	47
3	73	64881	62341	gnlIPIDle2684 56	unknown [Mycobacterium tuberculosis]	73	54
3	100	89800	91113	gil500705	Similar to Seryl-tRNA synthetase [Saccharomyces cerevisiae]	73	56
3	106	92803	93513	gnlIPIDle2436 81	ORF YGR248w [Saccharomyces cerevisiae]	73	63
4	4	3697	3512	gil562035	NADH dehydrogenase, subunit 5 [Acanthamoeba castellanii]	73	53
7	9	8519	8079	gil694092	emml gene product [Streptococcus pyogenes]	73	40

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

8	16	17562	17756	gil1500401	reverse gyrase [Methanococcus jannaschii]	73	40
14	3	4280	4438	gil520778	protein p23 [Borrelia burgdorferi]	73	55
19	9	7074	6742	gil1773311	NADH dehydrogenase [Ceanothus cuneatus]	73	36
25	3	2369	2587	gil1655790	CdsC [Borrelia burgdorferi]	73	64
78	2	176	619	gnl1PID1e201250	ORF-D gene product [Borrelia burgdorferi]	73	50
108	1	2	382	gil1573074	adhesin B precursor (fimA) [Haemophilus influenzae]	73	41
120	1	97	342	gil1978	heat shock protein 70 [Sus scrofa]	73	46
3	64	51644	54013	gil1574437	sporulation protein (spoIIIE) [Haemophilus influenzae]	72	51
5	6	2899	2654	gil212383	myosin heavy chain [Gallus gallus]	72	41
6	31	22140	21799	gil895748	putative cellobiose phosphotransferase enzyme II' [Bacillus subtilis]	72	46
8	8	8812	9600	gil1655859	Orf1 [Borrelia hermsii]	72	55
10	12	8579	8376	gil536681	ORF YBR257w [Saccharomyces cerevisiae]	72	36
45	2	1440	394	gil1699017	ErpB2 [Borrelia burgdorferi]	72	42
2	2	1342	2796	gil285623	pyruvate kinase [Bacillus stearothermophilus]	71	52
2	31	26272	24911	pirS58522IS58522	glycyl-tRNA synthetase - Thermus thermophilus	71	54
2	64	60156	58684	gil459009	similar to multifunctional aminoacyl-tRNA synthetase, especially to the prolyl-tRNA synthetase region [Caenorhabditis elegans]	71	48
3	66	55240	54275	gil217121	ORF1 [Synecococcus elongatus]	71	52
3	104	92345	92175	gil44228	secretion protein SecY (AA 1-482) [Mycoplasma capricolum]	71	42
5	43	25567	25734	gil213778	sodium-hydrogen exchange protein-beta [Oncorhynchus mykiss]	71	50
7	3	1179	2384	gil458216	ORF 1 [Borrelia burgdorferi]	71	60
20	4	2964	2392	gil458217	ORF 2 [Borrelia burgdorferi]	71	47
51	2	984	2066	gil1373144	ErpD [Borrelia burgdorferi]	71	41
54	1	251	883	gil145280	ORF1 [Escherichia coli]	71	40

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

2	54	51233	50538	spIQ06797IRL1_BACSU	50S RIBOSOMAL PROTEIN L1 (BL1).	70	48
2	116	114025	113744	gil1673757	(AE000012) Mycoplasma pneumoniae, phosphocarrier protein HPr; similar to GenBank Accession Number A49683, from M. capricolum [Mycoplasma pneumoniae]	70	41
3	4	1684	2220	gil153906	CheW protein [Salmonella typhimurium]	70	48
3	84	74775	73225	gnlPIDle283919	glycerol kinase [Sulfolobus solfataricus]	70	60
3	107	93500	93273	gil836815	cdc4 gene product which is essential for initiation of DNA replication in yeast [Saccharomyces cerevisiae]	70	35
4	1	926	123	gil167913	Thy1 protein [Dictyostelium discoideum]	70	51
4	47	35616	35807	gil48808	dciAE gene product [Bacillus subtilis]	70	58
4	65	48320	47976	gil1421734	ORF 5 [Borrelia burgdorferi]	70	48
6	23	16458	15904	gil1655860	Orf2 [Borrelia hermsii]	70	54
17	4	2940	3173	gil1255880	F01G12.6 gene product [Caenorhabditis elegans]	70	40
20	8	5470	5237	gil1236921	Var1p [Saccharomyces douglasii]	70	47
23	5	4173	3970	pirS16447IS16447	NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - wheat mitochondrion	70	47
36	3	1270	1653	gnlPIDle160437	orfD gene product [Borrelia burgdorferi]	70	50
2	69	65752	63860	gil151932	fructose enzyme II [Rhodobacter capsulatus]	69	42
3	114	99712	98150	gil1303856	Yqgl [Bacillus subtilis]	69	46
4	36	25614	24694	gil1663561	orf1; product unknown [Borrelia burgdorferi]	69	46
6	21	14584	14204	gil1616644	P30 [Borrelia burgdorferi]	69	47
12	12	7025	7258	gil150176	protein 69 [Mycoplasma hyorhinis]	69	38
12	14	8414	8587	gil13233	ND6 (AA 1 - 296) [Podospora anserina]	69	26
54	2	1332	2402	gnlPIDle158979	orfA gene product [Borrelia burgdorferi]	69	46
2	35	29769	30518	gil473817	'ORF' [Escherichia coli]	68	42
2	79	72330	72980	gil1498049	adenylate kinase [Paracoccus denitrificans]	68	37

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

2	104	104748	106385	gil1574032	hypothetical [Haemophilus influenzae]	68	42
3	78	68895	68287	gnlPIDle2551 17	hypothetical protein [Bacillus subtilis]	68	51
3	98	88992	86074	gil927711	D9461.18p; CAI: 0.15 [Saccharomyces cerevisiae]	68	52
3	111	96519	97364	gil1707057	coded for by C. elegans cDNA CEES55F; coded for by C. elegans cDNA yk84a1.3; coded for by C. elegans cDNA yk78g7.3; coded for by C. elegans cDNA yk168g9.5; coded for by C. elegans cDNA yk78g7.5; coded for by C. elegans cDNA yk84a1.5; strong s	68	52
4	56	40648	40046	gil458217	ORF 2 [Borrelia burgdorferi]	68	54
4	57	41916	40678	gil458216	ORF 1 [Borrelia burgdorferi]	68	57
6	24	17296	16520	gil1655859	Orf1 [Borrelia hermsii]	68	43
7	5	2894	3694	gil1655859	Orf1 [Borrelia hermsii]	68	48
29	6	3832	3254	gil458217	ORF 2 [Borrelia burgdorferi]	68	46
72	2	927	1133	gil577175	L8479.4 gene product [Saccharomyces cerevisiae]	68	44
2	57	52752	52558	gil1001264	50S ribosomal protein L33 [Synecocystis sp.]	67	56
3	65	54290	54051	gil710340	ribosomal protein S21 [Myxococcus xanthus]	67	49
3	79	69068	70114	gil460955	TagE [Vibrio cholerae]	67	38
3	81	70653	71150	gil467420	unknown [Bacillus subtilis]	67	42
3	110	94703	96502	gnlPIDle2676 07	alanyl-tRNA synthetase [Thermus aquaticus thermophilus]	67	51
4	42	30304	31941	bbsl161785	60 kDa antigen [Borrelia coriariae, C053, ATCC 4338, Peptide, 514 aa] [Borrelia coriariae]	67	49
12	6	3590	2967	gnlPIDle1604 37	orfD gene product [Borrelia burgdorferi]	67	41
12	9	5524	6276	gil1655859	Orf1 [Borrelia hermsii]	67	51
12	10	6611	6889	gnlPIDle8903	SERA protein [Plasmodium falciparum]	67	48
17	6	4995	5906	gil1752736	gene required for phosphorylation of oligosaccharides/ has high homology with YJR061w [Saccharomyces cerevisiae]	67	37
34	2	1221	1817	gnlPIDle1589	orfB gene product [Borrelia burgdorferi]	67	47

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

58	2	1347	796	83	ORF 2 [Borrelia burgdorferi]	67	52
2	33	28572	27751	gil340613	A 'c' was inserted after nt 369 (=nt 10459 in genomic sequence (M10126)) to correct -1 frameshift probably due to gel compression [Leishmania tarentolae]	66	40
2	73	69021	69908	gil153903	methyltransferase (cheR; EC 2.1.1.24) [Salmonella typhimurium]	66	42
2	93	93739	94524	gil45713	P. putida genes rpmH, rnpA, 9k, 60k, 50k, gidA, gidB, uncI and uncB [Pseudomonas putida]	66	41
3	9	6009	6902	gnlIPIDle2639 31	OrfD [Streptococcus pneumoniae]	66	47
4	28	20922	20665	gil471731	vacuolating cytotoxin homolog [Helicobacter pylori]	66	50
4	64	47985	47107	gil1421735	ORF 6 [Borrelia burgdorferi]	66	43
6	13	7227	8591	gil1591045	hypothetical protein (SP:P31466) [Methanococcus jannaschii]	66	48
34	4	2556	3161	gil458218	ORF 3 [Borrelia burgdorferi]	66	42
37	1	982	689	gil974334	non-receptor tyrosine kinase [Dictyostelium discoideum]	66	55
3	77	68191	66395	gil1651216	Pz-peptidase [Bacillus licheniformis]	65	47
3	123	105911	104070	gil1575784	DNA mismatch repair protein [Aquifex pyrophilus]	65	45
6	9	5726	7126	gil1591045	hypothetical protein (SP:P31466) [Methanococcus jannaschii]	65	49
8	9	9684	10325	gnlIPIDle2012 50	ORF-D gene product [Borrelia burgdorferi]	65	48
10	1	3	971	gil1373144	ErpD [Borrelia burgdorferi]	65	47
13	5	3956	3411	gil1209872	REV [Borrelia burgdorferi]	65	47
2	76	70509	71069	pirA00547IX YEBET	protein-glutamate methyltransferase (EC 3.1.1.61) - Salmonella typhimurium	64	45
3	61	48610	50838	gil1001335	soluble lytic transglycosylase [Synecocystis sp.]	64	42
4	5	3519	3773	gil1263021	M protein [Streptococcus pyogenes]	64	32
4	53	38288	37824	gil1373141	ORF-10 [Borrelia burgdorferi]	64	50

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

6	10	5985	5824	gil40271	delta-endotoxin CryIG protoxin [Bacillus thuringiensis]	64	30
7	7	7798	4499	gil1041785	rhostry protein [Plasmodium yoelii]	64	35
7	30	19738	19289	gil1209840	2.9-3 ORF-D [Borrelia burgdorferi]	64	46
11	3	1608	2339	gil1652934	hypothetical protein [Synecocystis sp.]	64	30
16	1	537	839	gnllPIDe276380	AARP1 protein [Plasmodium falciparum]	64	44
19	1	308	1177	gil1553115	P35 antigen protein [Borrelia burgdorferi]	64	35
42	3	1928	1788	gil1752736	gene required for phosphorylation of oligosaccharides/ has high homology with YJR061w [Saccharomyces cerevisiae]	64	41
142	1	589	2	gil162142	kinetoplast-associated protein [Trypanosoma cruzi]	64	52
2	3	2837	2592	gnllPIDe236274	ZK287.2 [Caenorhabditis elegans]	63	27
2	15	12750	11320	gil1652577	carboxyl-terminal protease [Synecocystis sp.]	63	49
2	32	27753	26266	spIP15189ISY E_RHIME	GLUTAMYL-TRNA SYNTHETASE (EC 6.1.1.17) (GLUTAMATE--TRNA LIGASE) (GLURS)	63	48
2	77	71067	72308	gil1041116	TRAB [Plasmid pPDI]	63	34
3	2	1056	58	gil1098641	Bts1p [Saccharomyces cerevisiae]	63	43
3	82	71398	71237	gil1339938	EC 1.1.99.5 [Mus musculus]	63	47
3	83	72845	71349	gil763191	glycerol 3 phosphate dehydrogenase [Saccharomyces cerevisiae]	63	37
3	85	75552	74773	gil142997	glycerol uptake facilitator [Bacillus subtilis]	63	45
7	6	3747	4304	gnllPIDe201250	ORF-D gene product [Borrelia burgdorferi]	63	40
7	38	24123	24956	gil467330	replicative DNA helicase [Bacillus subtilis]	63	40
11	5	4161	3853	gil1592217	bifunctional protein [Methanococcus jannaschii]	63	38
12	13	9558	7906	gil633167	adenine deaminase [Bacillus subtilis]	63	48
32	1	753	268	gil520783	unknown [Borrelia burgdorferi]	63	42
2	68	63866	62745	gil146722	phosphomannose isomerase [Escherichia coli]	62	45
2	75	69920	70573	gil145524	cheB peptide [Escherichia coli]	62	28

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

2	95	96334	95492	gil40031	spoOJ93 gene product [Bacillus subtilis]	62	36
3	67	57341	55212	gil1574144	single-stranded-DNA-specific exonuclease (recJ) [Haemophilus influenzae]	62	40
3	76	66414	65677	gil1477770	unknown [Helicobacter pylori]	62	37
6	1	1762	104	gil1072419	gicB gene product [Staphylococcus carnosus]	62	43
18	4	4431	5144	gil1591493	glutamine transport ATP-binding protein Q [Methanococcus jannaschii]	62	36
19	8	6743	6976	gil1513302	CigB [Dictyostelium discoideum]	62	56
20	6	4563	4378	bbs1144872	Fu=putative serine/threonine kinase [Drosophila melanogaster, Peptide Partial Mutant, 152 aa] [Drosophila melanogaster]	62	37
81	1	56	538	gnlPIDle1539 57	ORF-A gene product [Borrelia burgdorferi]	62	36
106	2	586	356	gil1151158	repeat organellar protein [Plasmodium chabaudi]	62	43
114	1	138	629	gnlPIDle1539 57	ORF-A gene product [Borrelia burgdorferi]	62	36
2	117	114352	114032	gil173128	ubiquitin-specific processing protease [Saccharomyces cerevisiae]	61	32
3	55	42737	44236	gil143999	dnaK homologue [Borrelia burgdorferi]	61	41
3	57	44821	46083	gil1653709	lipoprotein NlpD [Synecocystis sp.]	61	50
3	125	110052	109261	gil1303863	YqgP [Bacillus subtilis]	61	45
4	63	47119	46478	gil1421736	ORF 7 [Borrelia burgdorferi]	61	34
7	35	21496	22971	gil1655797	CdsJ [Borrelia burgdorferi]	61	44
8	7	8300	8872	gil458217	ORF 2 [Borrelia burgdorferi]	61	48
12	8	5006	5551	gil458217	ORF 2 [Borrelia burgdorferi]	61	50
14	10	9398	8652	gnlPIDle2012 50	ORF-D gene product [Borrelia burgdorferi]	61	44
15	12	6206	4377	gil836624	methyltransferase [Bacillus aneurinolyticus]	61	38
16	4	2449	2240	gil1066497	Similar to S. cerevisiae hypothetical protein Ykl012p (Swiss Prot. accession number P33203) and C. elegans hypothetical protein ZK1098.1 (Swiss Prot. accession number P34600) [Saccharomyces	61	38

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

29	4	2323	1853	gnlPIDle1604 37	cerevisiae] orfD gene product [Borrelia burgdorferi]	61	45
50	2	1374	1156	gnlPIDle2763 80	ARP1 protein [Plasmodium falciparum]	61	52
2	18	14371	15369	gil1573923	prolipoprotein diacylglycerol transferase (lgt) [Haemophilus influenzae]	60	57
3	118	101571	100690	gil1001260	hypothetical protein [Synecocystis sp.]	60	47
3	120	101692	102273	gil1399829	elongation factor P [Synecococcus PCC7942]	60	34
6	32	21869	22162	gil192960	L-type calcium channel alpha-1 [Mus musculus]	60	50
7	37	23373	24101	gil458217	ORF 2 [Borrelia burgdorferi]	60	40
8	11	13570	13851	gil1065989	(pos:5995..5997,aa:Met) [Bacillus subtilis]	60	47
14	5	5327	5091	gil147158	pfs [Escherichia coli]	60	51
15	7	3316	2984	gil153727	M protein [group G streptococcus]	60	36
27	3	2744	3772	pirS40422IS4 0422	hypothetical protein - Staphylococcus aureus	60	31
2	62	57446	58672	gil143002	proton glutamate symport protein [Bacillus caldotenax]	59	34
2	82	74989	74051	gil1651878	regulatory components of sensory transduction system [Synecocystis sp.]	59	38
2	89	92119	91322	gil467425	unknown [Bacillus subtilis]	59	38
2	92	93010	93663	pirA30191IA3 0191	hypothetical protein L - Bacillus subtilis (fragment)	59	39
2	118	115604	114315	gil39269	sigma factor (ntrA) (AA 1-502) [Azotobacter vinelandii]	59	35
4	41	29875	29210	gil1209831	lipoprotein [Borrelia burgdorferi]	59	34
6	4	3323	2058	gil624056	contains 4 ankyrin repeats; similar to D. melanogaster notch protein, Swiss-Prot Accession Number P07027 [Paramecium bursaria Chlorella virus 1]	59	37
6	25	17793	17257	gnlPIDle2012 48	ORF-B gene product [Borrelia burgdorferi]	59	43

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

6	33	22493	22125	gil153677	enzyme III [Streptococcus mutans]	59	36
10	9	6241	6026	gil1339977	skeletal myosin heavy chain [Thunnus thynnus]	59	40
19	6	5383	5970	gil160299	glutamic acid-rich protein [Plasmodium falciparum]	59	30
25	7	4008	3742	gil1055144	similar to galactoside 3(4)-L-fucosyltransferase [Caenorhabditis elegans]	59	37
59	1	835	8	gil1359436	Mag44 [Dermatophagoides farinae]	59	24
2	27	21414	20317	gil974332	NAD(P)H-dependent dihydroxyacetone-phosphate reductase [Bacillus subtilis]	58	41
3	5	2156	3049	gil1653618	hypothetical protein [Synecocystis sp.]	58	35
3	109	93822	94718	gil790935	fliG [Treponema denticola]	58	31
4	3	2423	3340	gil1553115	P35 antigen protein [Borrelia burgdorferi]	58	39
4	35	24696	24238	gil1663562	orf1; product unknown [Borrelia burgdorferi]	58	32
4	46	35509	34904	gnlPIDle264708	myosin heavy chain [Sus scrofa]	58	31
15	8	3683	3468	gil457336	Pv200 [Plasmodium vivax]	58	32
50	3	1941	1498	gnlPIDle220350	bcr2 gene product [Homo sapiens]	58	41
55	1	2322	247	gil1522636	M. jannaschii predicted coding region MJEC502 [Methanococcus jannaschii]	58	32
2	53	50563	50045	gil786163	Ribosomal Protein L10 [Bacillus subtilis]	57	29
3	117	100606	99710	gil1303855	YqgH [Bacillus subtilis]	57	30
6	38	26564	26232	gil1499632	M. jannaschii predicted coding region MJ0809 [Methanococcus jannaschii]	57	40
8	10	12350	13117	gil1553115	P35 antigen protein [Borrelia burgdorferi]	57	33
11	4	3183	2470	pinA456051A45605	mature-parasite-infected erythrocyte surface antigen MESA - Plasmodium falciparum	57	35
14	8	7117	7899	gil1553115	P35 antigen protein [Borrelia burgdorferi]	57	36
15	6	3027	2818	gnlPIDle261409	nuclear/mitotic apparatus protein [Xenopus laevis]	57	30
17	1	336	1178	gil473817	ORF [Escherichia coli]	57	31
20	2	1654	1064	gnlPIDle201250	ORF-D gene product [Borrelia burgdorferi]	57	34

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

2	111	113765	111636	gll148316	NaH-antiporter protein [Enterococcus hirae]	56	32
3	80	70112	70669	gll1372995	OrfH [Borrelia burgdorferi]	56	24
3	116	98976	99212	pirE22845IE22845	hypothetical protein 4 - Trypanosoma brucei mitochondrion (SGC6)	56	36
6	26	18732	17791	gll1655797	CdsJ [Borrelia burgdorferi]	56	41
7	21	14706	13510	gll1574247	H. influenzae predicted coding region HI1410 [Haemophilus influenzae]	56	32
11	8	6722	7087	gnlPIDle242897	aBIM [Lactococcus lactis]	56	28
53	7	2446	2018	gll1421737	ORF 8 [Borrelia burgdorferi]	56	38
61	2	712	1410	gll583161	albumin binding protein [unidentified]	56	35
2	6	3866	3573	gll290487	50S ribosomal subunit protein L28 [Escherichia coli]	55	37
2	14	11322	10585	gll1303811	YqeU [Bacillus subtilis]	55	33
2	34	28640	29782	gll558266	orf gene product [Wolinella succinogenes]	55	30
2	71	66669	67415	gll397486	endonuclease G [Bos taurus]	55	33
3	87	75924	76550	gll403984	deoxyguanosine kinase/deoxyadenosine kinase(I) subunit [Lactobacillus acidophilus]	55	38
4	66	48434	48958	gll1100900	70 kDa heat shock protein [Theileria parva]	55	32
140	1	322	68	gll15611	gene 17, tail fiber protein [Bacteriophage T7]	55	38
4	34	24244	23867	gll1663563	orfIII; product unknown [Borrelia burgdorferi]	54	31
5	9	5510	4179	gll1513238	ORFveg132; similar to Caenorhabditis elegans ORF F59B10.1 encoded by EMBL Accession Number Z49132 [Dictyostelium discoideum]	54	25
5	45	27187	25895	gnlPIDle261409	nuclear/mitotic apparatus protein [Xenopus laevis]	54	30
7	28	17905	18162	gll36501	C protein [Homo sapiens]	54	41
11	6	4415	5215	gll1707287	putative outer membrane protein [Borrelia burgdorferi]	54	25
19	2	1674	2501	gll392799	G5/D6 ORF [Dictyostelium discoideum]	54	25
29	5	3284	2532	gnlPIDle158980	orfC gene product [Borrelia burgdorferi]	54	33
31	3	3328	4137	pirS41649IS4	DNA polymerase - Plasmodium falciparum	54	28

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

						jannaschii			
4	9	6362	7153	gil1553115	P35 antigen protein [Borrelia burgdorferi]		51	26	
10	10	6603	7196	gnlPIDle2563 93	anti-P.falci parum antigenic polypeptide [Saimiri sciureus]		51	34	
11	12	10333	9422	pirA42771IA4 2771	reticulocyte-binding protein 1 - Plasmodium vivax		51	31	
19	7	5919	6179	gil173241	ZIP1 protein [Saccharomyces cerevisiae]		51	38	
23	1	3	287	gil1498320	cell wall-associated protease precursor [Bacillus subtilis]		51	25	
2	105	106383	107126	gil580905	B. subtilis genes rpmH, rnpA, 50kd, gidA and gidB [Bacillus subtilis]		50	32	
3	1	1	195	gnlPIDle2202 01	rps5 gene product [Plasmodium falciparum]		50	38	
3	62	50808	51653	gil882579	CG Site No. 29739 [Escherichia coli]		50	31	
3	119	100766	101014	gil1086864	T03G11.2 gene product [Caenorhabditis elegans]		50	39	
4	32	23555	22992	gil1663565	orfV; product unknown [Borrelia burgdorferi]		50	36	
5	8	4168	3470	gil49402	M1.1 protein [Streptococcus pyogenes]		50	27	
10	7	5190	4612	gnlPIDle1589 81	orfE gene product [Borrelia burgdorferi]		50	28	
11	2	1277	504	gil1553115	P35 antigen protein [Borrelia burgdorferi]		50	26	
13	3	1948	1634	gnlPIDle2682 43	p21 [Borrelia afzelii]		50	32	
92	3	582	941	gnlPIDle2012 50	ORF-D gene product [Borrelia burgdorferi]		50	40	
148	1	339	4	gnlPIDle2369 01	unknown [Saccharomyces cerevisiae]		50	34	
28	3	2001	2630	gil499325	STARP antigen [Plasmodium falciparum]		49	22	
3	10	6881	7180	gil156218	putative [Caenorhabditis elegans]		48	32	
3	75	65683	65066	gil1574476	dedA protein (dedA) [Haemophilus influenzae]		48	22	
3	112	97006	96743	gil915207	gastric mucin [Sus scrofa]		48	27	
7	23	14743	14970	gil172294	protein-tyrosine phosphatase [Saccharomyces cerevisiae]		48	33	

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

11	11	7980	9293	gil1046137	M. genitalium predicted coding region MG422 [Mycoplasma genitalium]	48	28
28	4	2628	2825	gil1591322	chorismate mutase subunit B [Methanococcus jannaschii]	48	30
2	8	5526	6677	gnlPIDle220245	frameshift [Plasmodium falciparum]	47	27
2	60	55075	55803	gil710551	ankyrin 3 [Mus musculus]	47	29
2	94	94515	95240	gil1592264	type I restriction enzyme [Methanococcus jannaschii]	47	34
4	11	9057	9941	gil1553115	P35 antigen protein [Borrelia burgdorferi]	47	32
2	12	9986	9471	gil1022328	Four tandem repeats of a DNA-binding domain known as the AT-hook are found at the carboxy terminus of CarD. This protein has been purified and found to bind in vitro to a promoter region [Myxococcus xanthus]	46	28
3	89	78904	77324	gil1573271	apolipoprotein N-acyltransferase (cute) [Haemophilus influenzae]	46	32
6	36	24361	25719	gil1592272	ribosomal protein S19 [Methanococcus jannaschii]	46	23
10	13	9895	8816	gil160299	glutamic acid-rich protein [Plasmodium falciparum]	46	21
13	4	3412	3648	gnlPIDle275506	C41G6.i [Caenorhabditis elegans]	46	37
138	1	632	15	gil157006	bicaudalD protein [Drosophila melanogaster]	46	23
3	124	109271	105909	gil1499043	M. jannaschii predicted coding region MJ0263 [Methanococcus jannaschii]	45	23
4	17	14212	15465	pirS30782IS30782	integrin homolog - yeast (Saccharomyces cerevisiae)	45	23
23	4	3950	4852	gnlPIDle236901	unknown [Saccharomyces cerevisiae]	45	27
92	1	258	4	gnlPIDle236901	unknown [Saccharomyces cerevisiae]	45	27
3	90	79020	81044	gnlPIDle236483	F54G8.4 [Caenorhabditis elegans]	44	31
12	7	4075	5019	gil1151158	repeat organellar protein [Plasmodium chabaudi]	44	26

Borrelia burgdorferi - Putative coding regions of novel proteins similar to known proteins

17	3	1735	2142	pirA42771A4 2771	reticulocyte-binding protein 1 - Plasmodium vivax	44	26
22	7	4179	2827	gil563812	XCAP-C [Xenopus laevis]	44	20
31	2	1682	2761	gil1438951	cutinase negative acting protein [Fusarium solani f. sp. pisi]	43	23
3	7	4086	5186	gil343962	VAR1 protein [Candida glabrata]	42	25
28	1	110	496	gil157804	laminin B2 chain [Drosophila melanogaster]	42	23
28	5	2889	3833	pirS30782IS3 0782	integrin homolog - yeast (Saccharomyces cerevisiae)	42	18
34	1	209	1234	gil1655797	CdsJ [Borrelia burgdorferi]	42	27
65	3	1035	1415	gil1654220	variable major protein 16 [Borrelia hermsii]	42	34
2	11	9544	8486	gnlPIDle1632 6	MURF2 protein (AA 1-348) [Crithidia fasciculata]	41	26
3	122	104072	103017	gil1151158	repeat organellar protein [Plasmodium chabaudi]	41	20
18	6	5122	6366	gil1591494	M. jannaschii predicted coding region MJ0797 [Methanococcus jannaschii]	40	20
6	6	4662	3964	gil600448	var1 protein (aa 1-339) [Candida utilis]	39	24
4	10	7637	8914	gil1293695	microfilament sheath protein SHP3 [Litomosoides sigmodontis]	37	19

TABLE 5. *Borrelia burgdorferi* - Coding regions containing to know proteins

Contig ID	Orf ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	HSP nt length
2	20	15372	17402	gbIM90084	<i>Borrelia burgdorferi</i> 22 kD antigen	100	786
2	21	16672	16310	gbIM90084	<i>Borrelia burgdorferi</i> 22 kD antigen	100	56
2	22	17362	17099	gbIM90084	<i>Borrelia burgdorferi</i> 22 kD antigen	100	264
2	23	17415	17876	embIX70826IB BLA7	<i>B. burgdorferi</i> gene for lipoprotein	100	57
2	24	18522	17923	embIX70826IB BLA7	<i>B. burgdorferi</i> gene for lipoprotein	100	600
2	25	18606	20009	embIX78708IB BYSCI	<i>B. burgdorferi</i> (ZS7) YSCI-like gene	100	1404
2	26	19981	20295	embIX78708IB BYSCI	<i>B. burgdorferi</i> (ZS7) YSCI-like gene	99	314
2	38	32899	32174	gbIU49938	<i>Borrelia burgdorferi</i> potential virulence gene cluster membrane proteins BmpC (bmpC) and BmpA (bmpA), BmpB protein (bmpB), putative protein 4, Mg ion transporter MgtE (mgtE), protein kinase C1 inhibitor PKCI (pkci) genes, complete cds	98	130
2	39	33315	32863	gbIU49938	<i>Borrelia burgdorferi</i> potential virulence gene cluster membrane proteins BmpC (bmpC) and BmpA (bmpA), BmpB protein (bmpB), putative protein 4, Mg ion transporter MgtE (mgtE), protein kinase C1 inhibitor PKCI (pkci) genes, complete cds	100	453
2	40	34718	33333	gbIU49938	<i>Borrelia burgdorferi</i> potential virulence gene cluster membrane proteins BmpC (bmpC) and BmpA (bmpA), BmpB protein (bmpB), putative protein 4, Mg ion transporter MgtE (mgtE), protein kinase C1 inhibitor PKCI (pkci) genes, complete cds	99	1386
2	41	36211	34751	gbIU49938	<i>Borrelia burgdorferi</i> potential virulence gene cluster membrane proteins BmpC (bmpC) and BmpA (bmpA), BmpB protein (bmpB), putative protein 4, Mg ion transporter MgtE (mgtE), protein kinase C1 inhibitor PKCI (pkci) genes, complete cds	99	1461
2	42	36899	36288	gbIL24194	<i>Borrelia burgdorferi</i> immunodominant antigen P39	99	606

Borrelia burgdorferi - Coding regions containing to know proteins

2	43	37335	36811	gbIL35050l	gene, complete cds	98	457
					Borrelia burgdorferi (clone pB46) membrane lipoprotein A (bmpA) gene, 3' end, membrane lipoprotein (bmpB) gene, 5' end		
2	44	38426	37401	gbIL24194l	Borrelia burgdorferi immunodominant antigen P39 gene, complete cds	99	1026
2	45	39595	38462	gbIU49938l	Borrelia burgdorferi potential virulence gene cluster membrane proteins BmpC (bmpC) and BmpA (bmpA), BmpB protein (bmpB), putative protein 4, Mg ion transporter MgtE (mgtE), protein kinase C1 inhibitor PKCI (pkci) genes, complete cds	99	1134
2	46	40947	39838	gbIU35450l	Borrelia burgdorferi membrane protein D (bmpD) gene, complete cds	99	1110
2	47	41461	40961	gbIU35450l	Borrelia burgdorferi membrane protein D (bmpD) gene, complete cds	100	82
2	49	46052	41901	gbIL48488l	Borrelia burgdorferi RNA polymerase beta subunit (rpoB) gene, complete cds, RNA polymerase beta' subunit (rpoC) gene, 5' end of cds	97	76
2	51	49535	46050	gbIL48488l	Borrelia burgdorferi RNA polymerase beta subunit (rpoB) gene, complete cds, RNA polymerase beta' subunit (rpoC) gene, 5' end of cds	98	2490
2	83	79470	74977	gbIU03396l	Borrelia burgdorferi B31 Ala-tRNA (alaT), Ile-tRNA (ileT), 16S rRNA, 23S rRNA (rrlA and rrlB), and 5S rRNA (rrfA and rrfB) genes, complete sequence	97	131
2	84	84351	84620	gbIM88330l	Borrelia burgdorferi 23S ribosomal RNA gene	100	270
2	85	86923	86066	gbIU03396l	Borrelia burgdorferi B31 Ala-tRNA (alaT), Ile-tRNA (ileT), 16S rRNA, 23S rRNA (rrlA and rrlB), and 5S rRNA (rrfA and rrfB) genes, complete sequence	95	386
2	86	87637	87041	gbIU03396l	Borrelia burgdorferi B31 Ala-tRNA (alaT), Ile-tRNA (ileT), 16S rRNA, 23S rRNA (rrlA and rrlB), and 5S rRNA (rrfA and rrfB) genes, complete sequence	99	209

Borrelia burgdorferi - Coding regions containing to know proteins

2	87	88424	88116	gblU033961	Borrelia burgdorferi B31 Ala-tRNA (alaT), Ile-tRNA (ileT), 16S rRNA, 23S rRNA (rrlA and rrlB), and 5S rRNA (rrfA and rrfB) genes, complete sequence	96	210
2	88	91249	90680	gblU033961	Borrelia burgdorferi B31 Ala-tRNA (alaT), Ile-tRNA (ileT), 16S rRNA, 23S rRNA (rrlA and rrlB), and 5S rRNA (rrfA and rrfB) genes, complete sequence	100	570
2	96	98846	96393	embIZ121651B BGYRAG	B burgdorferi gyrA gene encoding DNA gyrase subunit A (partial)	96	289
2	97	100759	98837	gblU045271	Borrelia burgdorferi 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds	98	904
2	98	100893	102389	gblU045271	Borrelia burgdorferi 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds	100	1497
2	99	102618	103787	gblU045271	Borrelia burgdorferi 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds	99	1170
2	100	103786	103607	gblU045271	Borrelia burgdorferi 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds	100	180
2	101	103866	104177	gblU045271	Borrelia burgdorferi 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds	100	312

Borrelia burgdorferi - Coding regions containing to know proteins

2	102	104254	104424	gblU045271	ribosomal protein L34 (rpmH) genes, complete cds	100	171
2	103	104393	104764	gblU045271	Borrelia burgdorferi 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds	100	148
3	11	7200	8597	gblU43739I	Borrelia burgdorferi 212 DNA gyrase b subunit (gyrB) and ribonuclease P protein component (rnpA) genes, partial cds, DnaA protein (dnaA), DNA polymerase III beta subunit (dnaN), and ribosomal protein L34 (rpmH) genes, complete cds	100	1185
3	12	8581	9666	embIX96685IB BCDG	Borrelia burgdorferi fesmid clone 31, complete sequence	99	912
3	13	9664	10767	embIX96685IB BCDG	B.burgdorferi cell division genes	99	1104
3	14	10826	10614	embIX96433IB BFTSWQA	B.burgdorferi ftsW, ftsQ & ftsA genes	100	213
3	15	10797	11546	embIX96433IB BFTSWQA	B.burgdorferi ftsW, ftsQ & ftsA genes	99	750
3	16	11519	12787	gblU43739I	Borrelia burgdorferi fesmid clone 31, complete sequence	100	1269
3	17	12785	14008	gblU43739I	Borrelia burgdorferi fesmid clone 31, complete sequence	100	1224
3	18	14006	14701	gblU43739I	Borrelia burgdorferi fesmid clone 31, complete sequence	100	696
3	19	14699	15655	gblL76303I	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flfEFGHI, flbABC genes, complete cds	98	712
3	20	15653	16213	gblU43739I	Borrelia burgdorferi fesmid clone 31, complete sequence	100	561
3	21	16149	17552	embIX96685IB BCDG	B.burgdorferi cell division genes	97	1404

Borrelia burgdorferi - Coding regions containing to know proteins

3	22	17550	17993	gblU43739l	Borrelia burgdorferi fesmid clone 31, complete sequence	100	444
3	23	17996	18475	gblU43739l	Borrelia burgdorferi fesmid clone 31, complete sequence	100	480
3	24	18445	18822	gblL76303l	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flieFGHI, flbABC genes, complete cds	100	378
3	25	18777	20546	gblL76303l	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flieFGHI, flbABC genes, complete cds	100	1770
3	26	20544	21596	gblU43739l	Borrelia burgdorferi fesmid clone 31, complete sequence	100	1053
3	27	21575	22531	gblU43739l	Borrelia burgdorferi fesmid clone 31, complete sequence	100	957
3	28	22529	23860	gblL76303l	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flieFGHI, flbABC genes, complete cds	99	1332
3	29	23836	24288	gblL76303l	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flieFGHI, flbABC genes, complete cds	99	453
3	30	24269	24898	gblU43739l	Borrelia burgdorferi fesmid clone 31, complete sequence	100	630
3	31	24864	26084	gblL76303l	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flieFGHI, flbABC genes, complete cds	99	1221
3	32	26095	26541	gblL76303l	Borrelia burgdorferi ftsA gene, 3' end of cds, ftsZ, orf230, smf, hslVU, flgBCE, flieFGHI, flbABC genes, complete cds	100	447
3	33	26525	27874	gblU43739l	Borrelia burgdorferi fesmid clone 31, complete sequence	100	1350
3	34	27891	28121	gblU43739l	Borrelia burgdorferi fesmid clone 31, complete sequence	100	231
3	35	28112	28900	gblU43739l	Borrelia burgdorferi fesmid clone 31, complete sequence	100	789

Borrelia burgdorferi - Coding regions containing to know proteins

3	36	28894	29682	gblU43739	Borrelia burgdorferi fesmid clone 31, complete sequence	100	789
3	37	29680	30267	gblU43739	Borrelia burgdorferi fesmid clone 31, complete sequence	100	588
3	38	30302	31363	gblU43739	Borrelia burgdorferi fesmid clone 31, complete sequence	100	1062
3	39	31403	31750	gblU43739	Borrelia burgdorferi fesmid clone 31, complete sequence	100	348
3	40	31728	32369	gblU43739	Borrelia burgdorferi fesmid clone 31, complete sequence	100	642
3	41	32356	33144	gblL75945	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flhL, flhM, flhZ, flagellar export apparatus (fliPQR, flhB), flhF, flbE genes	99	789
3	42	32642	32436	gblL75945	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flhL, flhM, flhZ, flagellar export apparatus (fliPQR, flhB), flhF, flbE genes	100	207
3	43	33129	33416	gblL75945	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flhL, flhM, flhZ, flagellar export apparatus (fliPQR, flhB), flhF, flbE genes	100	288
3	44	33425	34237	gblL75945	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flhL, flhM, flhZ, flagellar export apparatus (fliPQR, flhB), flhF, flbE genes	100	813
3	45	34320	34072	gblL75945	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flhL, flhM, flhZ, flagellar export apparatus (fliPQR, flhB), flhF, flbE genes	100	249
3	46	34222	35355	gblL75945	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flhL, flhM, flhZ, flagellar export apparatus (fliPQR, flhB), flhF, flbE genes	99	1134

Borrelia burgdorferi - Coding regions containing to know proteins

3	47	35349	37457	gblL75945I	Borrelia burgdorferi flagellar hook protein (flgE), flbD, flagellar motor apparatus (motAB), flhL, flhM, flhZ, flagellar export apparatus (fliPQR, flhB), flhF, flhE genes	100	2109
3	48	37455	38627	gblU43739I	Borrelia burgdorferi fmsmid clone 31, complete sequence	100	1173
3	49	38612	39526	gblU43739I	Borrelia burgdorferi fmsmid clone 31, complete sequence	99	816
3	50	39765	39421	gblU43739I	Borrelia burgdorferi fmsmid clone 31, complete sequence	100	345
3	51	39524	40012	gblU43739I	Borrelia burgdorferi fmsmid clone 31, complete sequence	100	489
3	52	39985	41919	gblU43739I	Borrelia burgdorferi fmsmid clone 31, complete sequence	100	1935
3	53	41928	42248	gblU43739I	Borrelia burgdorferi fmsmid clone 31, complete sequence	100	286
3	58	47505	46090	gblM28682I	B burgdorferi promoter element DNA	100	78
3	68	57596	60034	gblL77216I	Borrelia burgdorferi (strain B31) protease (lon) gene, complete cds	99	2439
3	69	62374	60023	gblL77216I	Borrelia burgdorferi (strain B31) protease (lon) gene, complete cds	100	274
3	92	82181	83026	gblU35673I	Borrelia burgdorferi OrfR gene, partial cds, and S20, Hbb, OrfH and Rho genes, complete cds	99	542
3	93	83032	83358	gblU35673I	Borrelia burgdorferi OrfR gene, partial cds, and S20, Hbb, OrfH and Rho genes, complete cds	100	327
3	94	83365	83697	gblU48651I	Borrelia burgdorferi PIG histone-like protein HBbu (hbb) gene, complete cds	100	327
3	95	83695	84105	gblU35673I	Borrelia burgdorferi OrfR gene, partial cds, and S20, Hbb, OrfH and Rho genes, complete cds	99	411
3	96	84167	85732	gblU35673I	Borrelia burgdorferi OrfR gene, partial cds, and S20, Hbb, OrfH and Rho genes, complete cds	99	1566
3	97	85778	86038	gblU35673I	Borrelia burgdorferi OrfR gene, partial cds, and S20, Hbb, OrfH and Rho genes, complete cds	100	106

Borrelia burgdorferi - Coding regions containing to know proteins

4	2	1935	1147	gblU611421	Borrelia burgdorferi outer membrane porin protein Oms28 precursor (oms28) gene, complete cds	99	789
4	12	10037	11002	gblU594871	Borrelia burgdorferi P35 antigen protein gene, and 7.5 kDa lipoprotein gene, complete cds	100	966
4	13	11365	11153	gblU598591	Borrelia burgdorferi strain B31 6.6 kDa lipoprotein gene, complete cds	100	213
4	14	11577	12230	gblU594871	Borrelia burgdorferi P35 antigen protein gene, and 7.5 kDa lipoprotein gene, complete cds	100	373
4	15	12578	13414	gblM852161	Borrelia burgdorferi 27kD protein antigen gene (p27), complete cds	78	370
4	16	13511	13753	gblU224511	Borrelia burgdorferi 49kb linear plasmid small 12kDa lipoprotein gene, complete cds	99	243
4	23	18668	17793	gblL314271	Borrelia burgdorferi (clone BbK2.1) phoA fusion protein gene, partial cds	100	169
4	49	36694	36347	gblU758671	Borrelia burgdorferi decorin binding protein B (DbpB) gene, complete cds	99	329
4	50	36351	36929	gblU758671	Borrelia burgdorferi decorin binding protein B (DbpB) gene, complete cds	99	564
4	51	36838	36692	gblU758671	Borrelia burgdorferi decorin binding protein B (DbpB) gene, complete cds	100	147
4	52	37001	37624	gblU758661	Borrelia burgdorferi decorin binding protein A (DbpA) gene, complete cds	93	533
4	55	40073	39318	gblU425991	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	69	731
4	58	43349	42447	gblL231371	Borrelia burgdorferi (27985CT2) OspA gene, 3' end and OspB gene, complete cds	99	903
4	59	44228	43347	embA040091A04009	B. burgdorferi OspA gene and 5'flanking region	100	882
4	60	44792	44403	gblL197021	Borrelia burgdorferi outer surface protein A (ospA) and outer surface protein B (ospB) genes, complete cds	88	370
4	61	45198	44758	gblL197021	Borrelia burgdorferi outer surface protein A (ospA) and outer surface protein B (ospB) genes, complete cds	89	375

Borrelia burgdorferi - Coding regions containing to know proteins

4	62	46440	45382	gblL197021	cds	Borrelia burgdorferi outer surface protein A (ospA) and outer surface protein B (ospB) genes, complete cds	85	622
4	67	49363	50622	gblL340161	cds	Borrelia burgdorferi (clone 8) S1 gene, complete cds	99	1260
4	68	50708	51580	gblL340171	cds	Borrelia burgdorferi (clone 8) S2 gene, complete cds	99	837
4	69	52203	51655	gblL314231	cds	Borrelia burgdorferi (clone BbK2.14) phoA fusion protein gene, partial cds	99	292
4	70	53018	52488	gblL411511	cds	Borrelia burgdorferi (clone 8) s3 gene, complete cds	99	297
5	1	535	71	gblU606421		Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	91	465
5	2	1526	546	gblU606421		Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	89	374
5	4	2395	2129	gblL314251		Borrelia burgdorferi (clone BbK3.168) phoA fusion protein gene, partial cds	98	135
5	11	6832	6542	gblS667081		{target sequence for detection of Lyme disease agent} [Borrelia burgdorferi, B31, 30-kb circular plasmid pIP87, Plasmid, 416 nt]	97	290
5	12	7422	6817	gblU449141		Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds	87	595
5	13	8167	7565	gblU449141		Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds	84	147
5	14	9408	8284	gblU449141		Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds	72	568
5	15	10122	9427	gblU306171		Borrelia burgdorferi BbK2.11 (bbk2.10), complete cds	93	560
5	16	10533	11324	gblU449121		Borrelia burgdorferi plasmid cp32-1, erpA and erpB genes, complete cds	93	790
5	17	11590	11330	gblU449131		Borrelia burgdorferi plasmid cp32-4, erpH gene, complete cds	95	261

Borrelia burgdorferi - Coding regions containing to know proteins

5	18	11761	11588	gblU42599	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	96	173
5	19	13256	11808	gblU42599	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	95	1431
5	20	14187	13636	embX87201B BBRGABCD	B. burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	100	552
5	21	14727	14185	embX87201B BBRGABCD	B. burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	100	511
5	22	15588	14788	embX87201B BBRGABCD	B. burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	99	801
5	23	16097	15519	embX87201B BBRGABCD	B. burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	98	579
5	24	17276	16158	embX87201B BBRGABCD	B. burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	94	1075
5	25	17558	18526	gblU45425	Borrelia burgdorferi 2.9-5 locus, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	86	927
5	26	19040	18564	gblU45422	Borrelia burgdorferi 2.9-2 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	89	379
5	27	19712	19116	gblU45421	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	85	596
5	28	20164	19775	gblU45421	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	97	390
5	29	20504	20121	gblU45426	Borrelia burgdorferi 2.9-6 locus, ORF-A-D genes, complete cds and REP+ gene, partial cds	99	384
5	30	20799	20446	gblU96714	Borrelia burgdorferi B31 BlyA (blyA) and BlyB (blyB) genes, complete cds	99	354
5	31	21006	20797	gblU45426	Borrelia burgdorferi 2.9-6 locus, ORF-A-D genes, complete cds and REP+ gene, partial cds	97	210
5	32	21903	21076	gblU96714	Borrelia burgdorferi B31 BlyA (blyA) and BlyB	95	440

Borrelia burgdorferi - Coding regions containing to know proteins

5	33	21470	21625	gblU96714I	(blyB) genes, complete cds			
					Borrelia burgdorferi B31 BlyA (blyA) and BlyB (blyB) genes, complete cds	94	151	
5	34	22518	22051	gblU45421I	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	90	467	
5	35	22806	22516	gblU45421I	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	93	286	
5	36	23082	22840	gblU45421I	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	95	242	
5	37	23397	23080	gblU45421I	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	96	317	
5	38	23768	23388	gblU45421I	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	95	381	
5	39	24331	23750	gblU45421I	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	90	495	
5	51	29986	29417	gblU60642I	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	97	300	
5	52	30414	29980	gblU60642I	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	99	435	
5	53	30803	30357	gblU60642I	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	97	447	
5	54	31204	30740	gblU60642I	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	96	465	
5	55	31775	31215	gblU60642I	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	98	374	
5	59	33577	32804	gblL31425I	Borrelia burgdorferi (clone BbK3.168) phoA fusion protein gene, partial cds	87	135	

Borrelia burgdorferi - Coding regions containing to know proteins

6	16	9678	9022	gb U01894	Borrelia burgdorferi B31 outer surface protein C (ospC) gene, complete cds	100	657
6	17	9836	11425	gb L25883	Borrelia burgdorferi 26 kb plasmid GMP synthetase (guaA) gene, complete cds	98	1590
6	18	11435	12664	gb U13372	Borrelia burgdorferi 26 kb plasmid IMP dehydrogenase (guaB) gene, partial cds	100	1212
6	19	12195	11686	gb U13372	Borrelia burgdorferi 26 kb plasmid IMP dehydrogenase (guaB) gene, partial cds	100	510
7	1	695	3	gb U85588	Borrelia burgdorferi transposase-like protein (tra) gene, partial cds	99	693
7	2	1081	677	gb U85588	Borrelia burgdorferi transposase-like protein (tra) gene, partial cds	98	375
7	39	25041	25847	gb U45423	Borrelia burgdorferi 2.9-3 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	77	437
8	2	1420	746	gb U45424	Borrelia burgdorferi 2.9-4 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	80	193
8	12	14287	14087	gb U84396	Borrelia burgdorferi 16 kb plasmid DNA fragment	95	140
8	17	18352	17876	gb U85588	Borrelia burgdorferi transposase-like protein (tra) gene, partial cds	96	362
9	1	2815	2507	gb U43414	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	309
9	2	3522	2767	gb U43414	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	99	756
9	3	5188	5862	gb U43414	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	675
9	4	6809	7255	gb U43414	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	447
9	5	8621	7467	gb U43414	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	1155
9	6	9079	8735	gb U43414	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	345

Borrelia burgdorferi - Coding regions containing to know proteins

9	7	10224	9214	gblU43414I	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	911
9	8	10370	10972	gblU43414I	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	603
9	9	11844	11107	gblU43414I	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	738
9	10	13299	13027	gblU43414I	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	99	273
9	11	13612	13241	gblU43414I	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	99	372
10	2	2164	1604	gblU45422I	Borrelia burgdorferi 2.9-2 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	78	278
10	3	2686	2886	gblU12332I	Borrelia burgdorferi 16 kb plasmid hypothetical protein gene, complete cds	97	143
13	1	3	842	gblU03641I	Borrelia burgdorferi lp21 circular plasmid, complete sequence	91	290
13	2	1525	983	gblL16625I	Borrelia burgdorferi exported neurotoxin-like protein gene, complete cds	99	531
13	6	4098	4901	gblU03641I	Borrelia burgdorferi lp21 circular plasmid, complete sequence	86	713
13	7	4691	4467	gblU03641I	Borrelia burgdorferi lp21 circular plasmid, complete sequence	88	224
13	8	6348	5041	gblU42599I	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	82	1202
13	9	6673	7788	embX87127IB BPBRGEA	B.burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	81	519
13	10	7786	8355	gblU03641I	Borrelia burgdorferi lp21 circular plasmid, complete sequence	78	414
13	11	8393	8968	gblU03641I	Borrelia burgdorferi lp21 circular plasmid, complete sequence	84	576
13	12	9290	9544	gblU03641I	Borrelia burgdorferi lp21 circular plasmid, complete sequence	91	210

Borrelia burgdorferi - Coding regions containing to know proteins

14	6	5768	6217	gblL31616I	Borrelia burgdorferi protein p23 gene, complete cds	89	396
14	7	6126	6671	gblL31616I	Borrelia burgdorferi protein p23 gene, complete cds	85	242
16	5	3660	2854	gblM97452I	Borrelia burgdorferi outer surface protein D (ospD) gene, complete cds	100	807
19	3	3136	3657	gblL41151I	Borrelia burgdorferi (clone 8) s3 gene, complete cds	77	267
21	1	849	4	gblU60963I	Borrelia burgdorferi plasmid cp32-1 PCR target site, partial sequence	95	296
21	2	1427	834	gblU44914I	Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds	93	594
21	3	2168	1581	gblU44914I	Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds	83	130
21	4	2946	2257	gblU80956I	Borrelia burgdorferi strain 297CH putative outer membrane protein (ospF) gene, complete cds	91	350
21	5	3794	2964	gblU44913I	Borrelia burgdorferi plasmid cp32-4, erpH gene, complete cds	100	401
21	6	4334	5143	gblU44913I	Borrelia burgdorferi plasmid cp32-4, erpH gene, complete cds	99	413
21	7	5362	5183	gblU44913I	Borrelia burgdorferi plasmid cp32-4, erpH gene, complete cds	91	180
21	8	5581	5360	gblU42599I	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	97	221
22	1	306	4	gblU45421I	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	95	303
22	2	664	317	gblU45421I	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	97	336
22	3	1230	658	gblU45421I	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	89	406
24	1	69	4058	gblU76406I	Borrelia burgdorferi putative vls recombination cassettes Vls2-Vls16b (vls) gene, complete	97	2305

Borrelia burgdorferi - Coding regions containing to know proteins.

24	2	4056	5108	gb U764061	sequence	Borrelia burgdorferi putative vls recombination cassettes Vls2-Vls16b (vls) gene, complete sequence	94	750
25	1	383	760	gb U434141	sequence	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	378
25	2	1333	1536	gb U434141	sequence	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	204
26	1	684	82	emb X87127 B BPBRGEA	sequence	B.burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	96	603
26	2	903	682	emb X87127 B BPBRGEA	sequence	B.burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	96	221
26	4	2181	2573	gb AF000270	sequence	Borrelia burgdorferi strain B31 2.9-like locus, OrfC, OrfD, Rev (rev), lipoprotein (LP), and 36 kDa-like orf2 genes, complete cds, and 36 kDa-like orf1 gene, partial cds	94	362
26	5	3073	2621	gb U454271	sequence	Borrelia burgdorferi 2.9-7 locus, ORF-A-D, REV, and lipoprotein (LPA and LPB) genes, complete cds	80	220
26	6	3745	3149	gb U454231	sequence	Borrelia burgdorferi 2.9-3 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	87	478
26	8	4663	4355	gb U454241	sequence	Borrelia burgdorferi 2.9-4 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	98	309
27	1	997	434	gb L316151	sequence	Borrelia burgdorferi (clone BbK2.5-6) unknown protein gene, complete cds	96	219
27	2	1395	2258	gb L316161	sequence	Borrelia burgdorferi protein p23 gene, complete cds	98	610
30	1	252	686	emb X87127 B BPBRGEA	sequence	B.burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	97	419
30	2	760	1545	emb X87127 B BPBRGEA	sequence	B.burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	100	786
30	3	1543	2157	emb X87127 B	sequence	B.burgdorferi repeated DNA element, 30.5 kb	100	615

Borrelia burgdorferi - Coding regions containing to know proteins

				BPBRGEA	circular plasmid copy		
30	4	2158	2802	embIX87127IB BPBRGEA	B.burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	100	645
30	5	3247	4230	gbIU42599I	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	95	976
33	1	450	995	gbIU72996I	Borrelia burgdorferi plasmid cp32-5, erpI gene, complete cds	100	546
33	2	1008	2159	gbIU78764I	Borrelia burgdorferi plasmid cp32-1, erpA and erpB2 genes, complete cds	100	1152
33	3	2253	2882	gbIU44914I	Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds	98	379
33	4	3050	3628	gbIU44914I	Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds	93	577
35	1	3	176	gbIU03396I	Borrelia burgdorferi B31 Ala-tRNA (alaT), Ile-tRNA (ileT), 16S rRNA, 23S rRNA (rrlA and rrlB), and 5S rRNA (rrfA and rrfB) genes, complete sequence	91	174
35	2	976	737	gbIM88330I	Borrelia burgdorferi 23S ribosomal RNA gene	100	240
36	1	1	525	embIX87201IB BBRGABCD	B.burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	77	159
38	1	672	28	gbIU44914I	Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds	92	571
38	2	850	653	gbIU42598I	Borrelia burgdorferi plasmid cp32-3, ErpG (erpG) and BapA (bapA) genes, complete cds	100	133
38	3	1516	983	embIX82409IB BOSPG	B.burgdorferi ospG and bapA genes	100	534
38	4	2200	1604	embIX82409IB BOSPG	B.burgdorferi ospG and bapA genes	100	597
38	5	2602	3132	gbIU42598I	Borrelia burgdorferi plasmid cp32-3, ErpG (erpG) and BapA (bapA) genes, complete cds	99	529
39	1	967	665	embIX87202IB BBRGBCDE	B.burgdorferi plasmid, orfA, B, C, D, E, & G genes, clone pOMB10	97	170
39	2	1505	957	embIX87202IB	B.burgdorferi plasmid, orfA, B, C, D, E, & G	89	176

Borrelia burgdorferi - Coding regions containing to know proteins

					BBRGBCDE	genes, clone pOMB10		
39	3	2353	1553	2353	gblU42599	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	91	137
39	4	2574	2284	2574	embIX87201IB BBRGABCD	B. burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	94	291
39	5	2874	2572	2874	embIX87127IB BPBRGEA	B. burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	91	284
39	6	3028	2861	3028	embIX87127IB BPBRGEA	B. burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	93	168
40	1	596	132	596	gblU76406	Borrelia burgdorferi putative vls recombination cassettes Vls2-Vls16b (vls) gene, complete sequence	100	465
40	2	1753	575	1753	gblU76406	Borrelia burgdorferi putative vls recombination cassettes Vls2-Vls16b (vls) gene, complete sequence	100	1179
40	3	3000	1732	3000	gblU76406	Borrelia burgdorferi putative vls recombination cassettes Vls2-Vls16b (vls) gene, complete sequence	99	1269
41	1	1	411	1	embIX87201IB BBRGABCD	B. burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	99	411
41	2	342	1127	342	embIX87201IB BBRGABCD	B. burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	80	785
41	3	1172	1747	1172	embIX87201IB BBRGABCD	B. burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	80	572
41	4	1745	2338	1745	gblU42599	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	65	571
42	2	1133	1384	1133	gblU85588	Borrelia burgdorferi transposase-like protein (tra) gene, partial cds	93	236
43	1	360	4	360	embIX87127IB BPBRGEA	B. burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	96	356
43	2	635	1741	635	gblAF000270	Borrelia burgdorferi strain B31 2.9-like locus, OrfC, OrfD, Rev (rev), lipoprotein (LP), and 36 kDa-like orf2 genes, complete cds, and 36 kDa-like	90	392

Borrelia burgdorferi - Coding regions containing to know proteins

43	3	2242	1784	gbU454231	orfI gene, partial cds		85	421
43	4	2860	2318	gbU454211	Borrelia burgdorferi 2.9-3 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds		95	259
44	1	1158	178	gbU606421	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds		89	374
44	3	2531	1761	gbL314251	Borrelia burgdorferi (clone BbK3.168) phoA fusion protein gene, partial cds		99	135
45	1	287	3	gbU787641	Borrelia burgdorferi plasmid cp32-1, erpA and erpB2 genes, complete cds		84	153
45	3	2037	1453	gbL139241	Borrelia burgdorferi outer surface protein E (OspE) gene, complete cds		90	386
45	4	2663	2893	gbU449121	Borrelia burgdorferi plasmid cp32-1, erpA and erpB genes, complete cds		90	230
46	1	174	338	gbU425991	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds		96	91
46	2	259	966	gbU425991	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds		100	692
46	3	964	1527	gbU425991	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds		100	564
46	4	1509	2111	gbU425991	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds		99	603
46	5	2537	2851	embIX872011B BBRGABCD	B.burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17		98	315
47	1	2	526	gbU454251	Borrelia burgdorferi 2.9-5 locus, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds		95	525
47	2	1245	724	gbU454241	Borrelia burgdorferi 2.9-4 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds		94	483

Borrelia burgdorferi - Coding regions containing to know proteins

47	3	1971	1321	gbU454241	Borrelia burgdorferi 2.9-4 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	89	651
48	1	363	25	gbU449141	Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds	87	327
48	2	412	1182	gbU729971	Borrelia burgdorferi plasmid cp32-6, erpK gene, complete cds	100	91
48	3	2047	1244	gbU729971	Borrelia burgdorferi plasmid cp32-6, erpK gene, complete cds	99	804
49	1	713	18	gbU764061	Borrelia burgdorferi putative vls recombination cassettes Vls2-Vls16b (vls) gene, complete sequence	99	606
49	2	2308	704	gbU764061	Borrelia burgdorferi putative vls recombination cassettes Vls2-Vls16b (vls) gene, complete sequence	98	1596
51	1	613	2	gbU425991	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	99	612
51	3	2203	2487	gbU449141	Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds	86	269
52	1	3	236	embIX872021B BBRGBCDE	B. burgdorferi plasmid, orfA, B, C, D, E, & G genes, clone pOMB10	94	146
52	2	179	319	embIX872011B BBRGABCD	B. burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	94	140
52	3	250	1050	gbU425991	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	86	146
52	6	1650	2201	gbU425991	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	81	422
53	1	93	581	gbIAF0002701	Borrelia burgdorferi strain B31 2.9-like locus, OrfC, OrfD, Rev (rev), lipoprotein (LP), and 36 kDa-like orf2 genes, complete cds, and 36 kDa-like orf1 gene, partial cds	99	489
53	2	883	719	gbIAF0002701	Borrelia burgdorferi strain B31 2.9-like locus, OrfC, OrfD, Rev (rev), lipoprotein (LP), and 36 kDa-like orf2 genes, complete cds, and 36 kDa-like	100	101

Borrelia burgdorferi - Coding regions containing to know proteins

53	3	1107	811	gb AF000270	orf1 gene, partial cds	100	289
53	4	1447	1064	gb AF000270	Borrelia burgdorferi strain B31 2.9-like locus, OrfC, OrfD, Rev (rev), lipoprotein (LP), and 36 kDa-like orf2 genes, complete cds, and 36 kDa-like orf1 gene, partial cds	96	381
53	5	1742	1380	gb U454271	Borrelia burgdorferi strain B31 2.9-like locus, OrfC, OrfD, Rev (rev), lipoprotein (LP), and 36 kDa-like orf2 genes, complete cds, and 36 kDa-like orf1 gene, partial cds	93	362
53	6	1949	1740	gb U454261	Borrelia burgdorferi 2.9-7 locus, ORF-A-D, REV, and lipoprotein (LPA and LPB) genes, complete cds	98	210
57	1	3	434	gb U454221	Borrelia burgdorferi 2.9-2 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	92	326
57	2	1580	471	gb AF000270	Borrelia burgdorferi strain B31 2.9-like locus, OrfC, OrfD, Rev (rev), lipoprotein (LP), and 36 kDa-like orf2 genes, complete cds, and 36 kDa-like orf1 gene, partial cds	98	362
57	3	1837	2109	emb X87127 B BPBRGEA	B.burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	84	246
58	4	1573	1800	gb L314251	Borrelia burgdorferi (clone BbK3.168) phoA fusion protein gene, partial cds	90	118
60	1	668	111	emb X87201 B BBRGABCD	B.burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	75	519
60	2	1479	694	emb X87127 B BPBRGEA	B.burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	72	786
60	3	1907	1410	emb X87201 B BBRGABCD	B.burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	95	498
62	1	284	3	emb X87201 B BBRGABCD	B.burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	79	260
62	2	878	282	emb X87202 B	B.burgdorferi plasmid, orfA, B, C, D, E, & G	74	501

Borrelia burgdorferi - Coding regions containing to know proteins

				BBRGBCDE	genes, clone pOMB10		
62	3	1704	910	gbIU425991	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	78	351
64	1	563	54	gbIU434141	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	510
64	2	1320	1117	gbIU434141	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	204
66	1	647	75	gbIU606421	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	93	300
66	2	1075	641	gbIU606421	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	96	435
66	3	1530	1018	gbIU606421	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	94	440
70	1	3	275	gbIU967141	Borrelia burgdorferi B31 BlyA (blyA) and BlyB (blyB) genes, complete cds	98	207
70	2	217	600	gbIU454261	Borrelia burgdorferi 2.9-6 locus, ORF-A-D genes, complete cds and REP+ gene, partial cds	99	384
70	3	557	946	gbIU454231	Borrelia burgdorferi 2.9-3 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	98	390
70	4	1424	1083	gbIAF0002701	Borrelia burgdorferi strain B31 2.9-like locus, OrfC, OrfD, Rev (rev), lipoprotein (LP), and 36 kDa-like orf2 genes, complete cds, and 36 kDa-like orf1 gene, partial cds	99	342
75	1	2	925	gbIU606421	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	98	374
75	2	936	1328	gbIU606421	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	96	393
76	1	464	12	gbIU454221	Borrelia burgdorferi 2.9-2 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	85	281
76	2	1256	540	gbIU454251	Borrelia burgdorferi 2.9-5 locus, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete	91	552

Borrelia burgdorferi - Coding regions containing to know proteins

77	1	433		2	gblU45422l		cds		90	379
								Borrelia burgdorferi 2.9-2 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds		
77	2	1159		509	gblU45424l			Borrelia burgdorferi 2.9-4 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	97	651
81	2	657		1034	gblU43414l			Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	80	255
83	1	3		1202	gblU76406l			Borrelia burgdorferi putative vls recombination cassettes Vls2-Vls16b (vls) gene, complete sequence	99	1198
85	1	1		360	gblU45421l			Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	91	347
85	2	358		1008	gblU96714l			Borrelia burgdorferi B31 BlyA (blyA) and BlyB (blyB) genes, complete cds	84	440
85	3	791		636	gblU96714l			Borrelia burgdorferi B31 BlyA (blyA) and BlyB (blyB) genes, complete cds	80	151
86	1	891		289	gblU45422l			Borrelia burgdorferi 2.9-2 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	86	486
86	2	1151		954	gblU45427l			Borrelia burgdorferi 2.9-7 locus, ORF-A-D, REV, and lipoprotein (LPA and LPB) genes, complete cds	97	148
88	1	137		3	gblU60642l			Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	97	135
88	2	325		131	gblU60642l			Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	98	195
88	3	565		323	gblU60642l			Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	98	243
88	4	954		508	gblU60642l			Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	98	447

Borrelia burgdorferi - Coding regions containing to know proteins

88	5	1091	891	gblU60640l	Borrelia burgdorferi plasmid cp32-2, sequence at position 5kb	98	201
91	1	927	34	gblU45422l	Borrelia burgdorferi 2.9-2 locus, ORF-C gene, partial cds, ORF-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	97	313
93	1	162	578	gblU45421l	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	89	331
93	2	572	940	gblU45421l	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	96	368
94	1	3	245	gblU45425l	Borrelia burgdorferi 2.9-5 locus, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	97	243
94	2	749	282	gblAF000270l	Borrelia burgdorferi strain B31 2.9-like locus, OrfC, OrfD, Rev (rev), lipoprotein (LP), and 36 kDa-like orf2 genes, complete cds, and 36 kDa-like orf1 gene, partial cds	90	458
97	1	506	3	gblU44914l	Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds	94	472
98	1	827	264	gblU42599l	Borrelia burgdorferi plasmid cp18, OspE (ospE) gene, partial cds	70	380
99	1	175	408	gblU43414l	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	234
99	2	329	757	gblU43414l	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	98	220
101	1	207	440	gblU43414l	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	234
101	2	361	837	gblU43414l	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	99	477
102	1	3	911	gblU76406l	Borrelia burgdorferi putative vls recombination cassettes Vls2-Vls16b (vls) gene, complete sequence	99	889
104	1	388	242	gblU45426l	Borrelia burgdorferi 2.9-6 locus, ORF-A-D genes,	100	146

Borrelia burgdorferi - Coding regions containing to know proteins

104	2	595	386	gblU96714l	complete cds and REP+ gene, partial cds	100	210
107	1	2	811	gblU45425l	Borrelia burgdorferi B31 BlyA (blyA) and BlyB (blyB) genes, complete cds	95	789
109	1	264	4	gblL31616l	Borrelia burgdorferi 2.9-5 locus, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	86	201
109	2	598	173	gblL31616l	Borrelia burgdorferi protein p23 gene, complete cds	93	396
109	3	807	580	gblL31615l	Borrelia burgdorferi (clone BbK2.5-6) unknown protein gene, complete cds	99	228
110	1	1	456	gblU45421l	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	95	456
110	2	450	761	gblU45421l	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	93	310
111	1	787	215	gblU45421l	Borrelia burgdorferi 2.9-1 locus, ORF 5-8, ORF-A-D, REP+, REP-, and lipoprotein (LP) genes, complete cds	89	405
119	1	653	84	gblU60642l	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	98	300
121	1	719	123	gblU60642l	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	98	374
122	1	403	2	gblU44914l	Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds	85	391
128	1	175	408	gblU43414l	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	234
128	2	329	700	gblU43414l	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	99	356
129	1	458	697	emblX87201lB BBRGABCD	B.burgdorferi plasmid, orfA, B, C, D, E, & F genes, clone pOMB14 and pOMB17	100	238
132	1	234	467	gblU43414l	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	100	234

Borrelia burgdorferi - Coding regions containing to know proteins

132	2	388	660	gblU434141	Borrelia burgdorferi linear plasmid lp16 DNA, complete sequence	99	171
133	1	3	560	emblX871271B BPBRGEA	B.burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	80	243
134	1	339	4	emblX872021B BBRGBCDE	B.burgdorferi plasmid, orfA, B, C, D, E, & G genes, clone pOMB10	78	331
141	1	554	33	gblU967141	Borrelia burgdorferi B31 BlyA (blyA) and BlyB (blyB) genes, complete cds	100	513
141	2	124	276	gblU967141	Borrelia burgdorferi B31 BlyA (blyA) and BlyB (blyB) genes, complete cds	100	153
143	1	67	498	gblU425981	Borrelia burgdorferi plasmid cp32-3, ErpG (erpG) and BapA (bapA) genes, complete cds	98	432
144	1	497	3	emblX871271B BPBRGEA	B.burgdorferi repeated DNA element, 30.5 kb circular plasmid copy	94	495
146	1	193	2	gblU036411	Borrelia burgdorferi lp21 circular plasmid, complete sequence	86	144
147	1	3	542	gblU606421	Borrelia burgdorferi plasmid cp32-4, sequence at position 4-6kb	88	296
153	1	352	2	gblM968471	Borrelia burgdorferi GrpE protein homologue gene, DnaK protein homologue gene, and DnaJ protein homologue gene, complete cds's	97	351

TABLE 6.

Borrelia burgdorferi - Putative coding regions of novel proteins not similar to know proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
2	4	2730	3554
2	5	3559	3410
2	7	5464	3869
2	13	10502	9999
2	17	13800	13576
2	19	15368	15204
2	28	21155	21400
2	50	41944	42186
2	58	53786	52911
2	59	54816	53773
2	61	57393	55813
2	63	57882	57682
2	65	60898	60203
2	66	61441	62070
2	67	62078	62692
2	70	65896	66540
2	74	70203	69910
2	78	71818	71399
2	80	72956	74032
2	81	73515	73267
2	90	92181	92525
2	91	92968	92555
2	108	109872	110057
2	112	112408	112812
2	113	112858	113037
2	114	113035	113460
2	115	113506	113724
2	119	114325	114852
3	6	3279	4079
3	8	5156	6019
3	54	42256	42789
3	59	47264	47506
3	60	47673	48692
3	63	51475	51026
3	70	60330	60575
3	71	61050	61349
3	72	61347	61670
3	74	63917	64303
3	86	75347	75532
3	88	76593	77384
3	99	89769	89005
3	102	91278	91661
3	103	92137	92463
3	105	92423	92785
3	108	93467	93886
3	115	98262	98681

Borrelia burgdorferi - Putative coding regions of novel proteins not similar to know proteins

3	121	102227	102904
3	126	111308	110055
4	6	3751	4179
4	7	4218	5042
4	19	16115	15516
4	20	17028	16075
4	21	17379	17092
4	22	17735	17397
4	24	19243	18785
4	25	18942	19196
4	26	20677	19259
4	27	19431	19751
4	29	21376	20876
4	30	21899	21423
4	31	22918	21845
4	33	23951	23553
4	37	26253	25627
4	38	26991	26332
4	39	28181	26931
4	40	29175	28522
4	43	30605	30342
4	45	34906	33548
4	48	35750	35932
5	3	2102	1527
5	5	2656	2393
5	7	3460	2900
5	10	6544	5645
5	40	25278	24322
5	41	25235	25600
5	42	25665	25276
5	44	25881	25663
5	47	27883	27410
5	48	28351	27881
5	49	29028	28324
5	50	29454	29026
5	56	32199	31666
5	57	32571	32200
5	58	32826	32569
5	60	32913	33245
5	61	33766	33575
5	62	34173	33742
5	64	35514	34861
6	2	954	1181
6	3	1590	1763
6	5	3400	3954
6	7	4691	5218
6	8	5187	5699

Borrelia burgdorferi - Putative coding regions of novel proteins not similar to know proteins

6	11	6498	5983
6	12	6975	6727
6	14	7978	7448
6	15	8479	7976
6	22	15106	15636
6	27	19999	18842
6	28	20036	20668
6	29	21814	20690
6	30	20949	21269
6	35	24136	23630
6	37	25697	26248
7	8	8100	7792
7	10	8145	8288
7	11	9374	8517
7	12	9771	9325
7	13	9652	10185
7	14	10163	9765
7	15	10517	10173
7	16	11363	10524
7	17	11904	11392
7	18	12495	11902
7	19	13516	12473
7	20	12807	13154
7	22	15149	14697
7	24	15855	15046
7	25	15503	15826
7	26	16638	15853
7	27	19344	16636
7	31	19473	19727
7	32	20067	19675
7	33	20762	20049
7	34	21136	20738
7	36	22975	23406
7	40	26667	25870
8	3	2907	4118
8	5	5898	6059
8	6	7399	8313
8	13	15645	15899
8	14	17281	16331
8	15	16905	17111
10	4	3211	3684
10	6	3857	4456
10	8	5982	5599
10	11	8038	7802
10	14	10255	10100
11	7	5688	5828
11	9	7248	7685

Borrelia burgdorferi - Putative coding regions of novel proteins not similar to know proteins

11	10	7672	8028
11	13	9642	10154
12	1	101	370
12	2	982	680
12	3	1390	1115
12	4	1528	1388
12	5	1913	1431
12	11	7308	6616
14	2	3588	3328
14	4	4657	4815
14	9	7981	8511
15	1	1	327
15	2	325	1077
15	3	1478	657
15	4	2360	1758
15	5	2839	2507
15	9	3922	3743
15	10	4145	3900
15	11	4112	4270
15	13	7677	6127
15	14	7852	7709
15	15	8052	7825
15	16	8222	7857
16	2	1733	1936
16	3	1905	2063
16	6	5212	4220
16	7	8903	8505
17	2	1500	1709
17	5	4097	4660
17	7	6344	6189
18	1	1635	2465
18	2	2509	3306
18	3	3332	4390
18	5	4933	4727
18	7	6353	7084
18	8	7098	7625
20	7	4700	4557
22	4	2175	1228
22	5	2132	2314
22	6	2829	2173
22	8	3254	3601
22	9	4408	4169
22	10	4875	4402
22	11	5343	4873
23	2	2283	1537
23	3	3564	2617
25	6	3677	4147

Borrelia burgdorferi - Putative coding regions of novel proteins not similar to know proteins

26	7	4251	3889
28	2	732	1739
29	3	310	885
31	1	28	195
32	3	935	1603
32	4	1637	2332
37	2	1379	1059
42	4	2708	2388
44	2	1734	1159
44	4	2942	2532
47	4	2336	2115
50	1	908	120
52	4	674	501
56	1	152	1465
56	2	611	459
56	3	1479	2150
58	3	1691	1329
58	5	1867	2046
59	2	2018	1044
61	1	1	657
61	3	1389	1907
62	4	1115	1345
63	1	663	325
63	2	769	446
63	3	1759	1013
65	1	472	903
65	2	901	1236
67	1	387	4
67	2	979	401
67	3	1482	961
68	2	451	612
69	3	840	574
71	1	363	4
72	1	586	933
73	1	300	4
73	2	824	279
73	3	1396	1145
79	1	22	1119
82	1	701	303
82	2	1188	775
84	1	331	134
84	2	983	348
87	1	277	2
87	2	1136	267
96	1	434	57
96	2	748	557
97	2	976	659

Borrelia burgdorferi - Putative coding regions of novel proteins not similar to know proteins

103	1	301	2
103	2	886	299
105	1	36	509
106	1	425	3
106	3	761	600
112	1	416	799
113	1	685	59
118	1	1	489
118	2	487	753
120	2	299	691
124	1	1	630
127	1	702	322
135	1	287	3
135	2	649	407
136	1	1	645
140	2	619	332
145	1	1	480

(1) GENERAL INFORMATION:

(i) APPLICANT: Human Genome Sciences, Inc. et al.

(ii) TITLE OF INVENTION: *Borrelia burgdorferi* Polynucleotides and Sequences

(iii) NUMBER OF SEQUENCES: 155

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Human Genome Sciences, Inc.

(B) STREET: 9410 Key West Avenue

(C) CITY: Rockville

(D) STATE: Maryland

(E) COUNTRY: USA

(F) ZIP: 20850

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

(B) COMPUTER: HP Vectra 486/33

(C) OPERATING SYSTEM: MSDOS version 6.2

(D) SOFTWARE: ASCII Text

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE: Herewith

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

157

- (A) NAME: Brookes, A. Anders
(B) REGISTRATION NUMBER: 36,373
(C) REFERENCE/DOCKET NUMBER: PB370PCT

(vi) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (301) 309-8504
(B) TELEFAX: (301) 309-8512

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 910715 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATATAATTTT TAATTAGTAT AGAATATGTT AAACTTTACC CTTGAATTTT TCTACTCTAT	60
TTGTATATTC TATAGAAAAA ACGATTAGAA TTAAACAAAG CCATAACTGA ACCAACGGTA	120
ATTAGTAGAT AAAGGGATCA AAATATTTTT TATTGCAGCA AGAATACCTT GGTATATTAG	180
AAAAACCAAA AGTCATAGTC AAATCATCTT TTGATAACAA TCCCCAAATC TATAATTTAT	240
TATGAAATTA ATTGCTCCCT TGAAAAGATT AGTTTTTAAA ACTACAAGAC TACTATCAAT	300
CACTATCAGA TAGATTAAAA CAACCTTTAC AAGAAAAAAA TCTTACTACT ATTTTATTGT	360
AAATGTATTA TAAAATAAGT TCATGCAAAA ACTTACAATT TTTCACAACA AACTACAATA	420
AAATCATGTA AACAAACAAT TTCTTTGAAA ATTAAGCAAA TTTATAAATA TAAATTATAA	480
AGATATATAT TTTTATATGA TCAATAATAA AAATTAATAG GATACTTATT TGGAAAAATT	540
ATTGAAAAAA CAATAAGCAT GAATTGCCAC AATAAGCTAA TTGTCACTTA ATAATTCTTG	600
TTTACTAGAC CACATTAGTA TAAACTCAA TATTGGCTAC TATAATATAG GGGCTTTATA	660
CGCCACATGT TTAATGATAA CATAAGAAAA TATTGCAATA ATAAAAAGAT TGAAATATCT	720
TTATTAGAAA AGAATCTCGA TAATTTAGAA AACAGAATAA AAATCATAAC TAATAAATAT	780
AACGTTGAAA AAAATATATT CAAACTTTAA CTATACAATT AATTACACCT TAAAAATGCG	840
TTACATAAAA ATTAAGGACT ACTATAAATA GAAAACACCA CATAACCTAC AGACTCTAAA	900
GGAATAATTA AATCCTCATA TTTCAGTTCT CCAAAAGTTT AAATAGGGGC CTTTTACTTT	960

TCTTGATTAG CATATACATT ATTAAAGGCA TCTTCTTGGG CACTATCCTA AACTTTTTTTA	1020
CATTATTATT ATTTTATTCT TTATTATTAC AAGATAATTC AAGAATCTAG ATTACAAGAT	1080
ATCAATCCTG CCATTAGTAG TTCAATAAAA CATTTAGAAT ATTTATACAT TATTTAATGT	1140
ATTTTTTTCA TTTTGTGAAAT AATATTGTTA TAACTTAACT TAATAAGATA TTTGATTTCT	1200
TCAACTTGAG AATCCGATGT ACATAGAATC TGAACATCTC CTCTGCCCCA TTTGCCAATA	1260
TTCTTAATAT ATCTAGTAAA ACCCTCTTTT AAAATTATTT GATCTAGAGC AACAGTAATA	1320
GTAATATTAA TTTTATTTAC CCCAGGTCTA AAGCTAAAAT CTACAAAATA TCCGCCCTGT	1380
ACTTTAAATC CTGTATAGCA CTGTGTTTCA ACTTTCTCAA TTTTATTAAA ATTTAAAACA	1440
AAAATAAAAT CTTCTAATTC TTTATATATT GCTTTCATAT CGGAATTTAA TTTTCAAAT	1500
TTTTTTAAAT TTTCGGTTTT AATATTATTA TCTTTTATAC CAGAATCTGT GTCATCTTCT	1560
ATGTCACTTT TCTTGCTGTT TACTAATACA TCGCTTTTTT TTTTATCAAA AACATACTA	1620
AAAATATTTT TAATAATATC ATTAAATATT TTATCTGAAT ATGTTTTTTT AAAACCAATT	1680
TTAGCTTTAA AAAATCAAG CAAATCAACA CTGGATTTT TTGTTTCCTT TTTTAAATAA	1740
GCTGAAAATT TGTCTGTATA TTTTTTTTCT AATGCAAAAG ATCTAGCCTC TTCAACATTC	1800
AAAGAATTTT TAGAAAACTT TTTAAGATAT TCAAAATCCT TAGATGTTAA TTTTCTAAA	1860
TTAACAACCA TAAAAGGCTC ATTGTCTAAC AAATTATCTT TATCTAGGTC AGTATAGAAT	1920
CTATATTCTA TGCCATCTGT TAATATACCA AATTCAACTC TCTTTGCTTG AGAACGAATA	1980
TTTTCAAAAT AAGGTTTTAA TTGCTTTAGA TGATTTTCAA GCTTTTCCCT GCTATTATGA	2040
TATTTGGCCT CTATTAAAAAT AGTGGGTCT TCATCCTTTT TTGTTGGATA AATAACATAA	2100
TCAACCCTTT TTAGTCCATC TTTAAGAATA TCTGCCTTCT CTTCAACTTT AACAATTGAA	2160
ATATCAGTAT GATCATAGCC CATCGCATCT AAAAATGGAT CAATAAGATT TTGTCTTGTT	2220
TGTGCTTCAT TTTCAATAAG ATCCTTATCC TTTTGAATTT TTCTACTTAC AGCTTTTATT	2280
GAATTTTCAA AATTTATATC TTTGTATTCA TTTGGCATAA TTATATTTTA CCAATAAAAT	2340
TAAAAATTAA TAATTCTAAA AATAAATTC CAAAATGTTG TCTATTTTAA ACTCTTAACT	2400
GATACCTTAA TTCTTTTTTC TACCTAATTT TTTAGTTTAA AATCTTATTT TTTAATTTTA	2460
TTATTTTTTC CTTACCTTAT TTATACTAAA ATTTTTAGTA TTTAGCGAAT AATTTTCATA	2520
TCCTTTTATT AAAGACAAAA TATGATTTTC TCTTTTTTGT TTTTAAATAC CTTAAATCA	2580
CTAAGCAAAG TAATAAGTC TTCTTTGGTT AATGAATAAA AGACTAGCTA TAATAAAATT	2640
ATTTTATTTT TCTTTACTAA ATTCAAAATG CTCTAAATAA AGCAAATTAG AGAAATTCAA	2700

AGGATCATTT	TTAGCTATTA	GCAGAGAAGT	GTTTTTTACC	AAAGTTAGAC	ATAATGAACT	2760
AGCCAAAATT	TCTTCTTTGG	GTTGAGGCAT	TGGACATTGA	CAAAGAAATG	ATTTTACAAT	2820
GTCGGTATTT	TAAAACAAAT	CTTCTAATCA	TAAAATCAAA	TACAGTGCAT	TGAAAATAGA	2880
TATAATAAAC	AATTTTTTTAT	AAAAAGATAT	TGGTATTTTC	TCACAATTCA	TATCTATTTT	2940
ATAGAAACAC	AATAATAATT	TTTAGGAGAT	AAAGTGCTAA	TCATGGTTCT	TTCATTTGTA	3000
TTGCTTGCAA	TTCTTCTATA	AAATATTCTT	TCATTTGGGT	ACTGATCATC	TTTAGTTAAG	3060
ATTTTTTCTA	AATCTTCTTT	ATATCCTATC	CATAAAAGCT	TATAACCTTC	TTTTACATAA	3120
TCATAAGTAA	AAAATCTTAA	ATTAAATTGA	TAGATATTAG	CCCCAGAATA	AAGAAATATA	3180
AAGTTTTTCAT	TATTATATTC	CTTTAATAAA	GATTTGCGAT	TCTTTATACT	TGGATCTGGC	3240
CCTTTTTTAA	AATTAATATC	TTCTTTACTA	AGAATACTAA	ATGAACTAAA	TATTTTGTTT	3300
AATTTGGCCC	ATGTTTAATT	CAATTCCTTT	ATAAGGATTT	TCTTTGCAGT	CTTTTAAGTC	3360
TCTAGTTATT	CCTTAATAAT	ATTATCACTA	CTTTGAATAA	CAAATTTTGC	TTTAAATTTT	3420
AATGTAAAAG	TTTATTACTA	CGAGGAAATA	TCGCAAATTT	AAACTTGAA	TGCATATCTT	3480
AAAACCTTTT	TTTGTTTTCA	AACTGATAAA	TAAGTTAAGT	TTATAATTAC	TAAATATATG	3540
CTTCTTAGC	AAGCTAAGAC	CAAATATCAC	AATAGAAGTA	ATTCTCAATA	AACAAAATAC	3600
AAAAAGTAGT	TATCATATCG	TCTTTAACCT	TAAATAAGGT	TGCTATAAAC	AACCAAGATA	3660
TTTAATTTCT	TTTAAAACCC	TTATTCAATC	TTTTTAAGCA	TAGGATCTTA	TAATTATAAG	3720
AATATAATTT	TATTTACATC	TCTATATTAA	TAGAAAGATG	CAAATATGTG	ATCAAATTGT	3780
TATTTTGTGA	ATATGGAATA	GTCCTTTATA	GGGACGCTTA	ATGCTCTATA	CTTAAGATTG	3840
GAATTCTCTA	TGAAAATATA	TACTCGCTAC	CCATGTAAAG	CTGACTTATT	TTAGCACGTA	3900
TCGCTTAAAC	AATTATATTT	ATATTATCTT	TTATAAAGTT	AATTTTTTCT	TGTAGATTAT	3960
TTTTTAATAA	AAAAGGCACA	AATTACCACA	ACAAGTTCCA	GTATAAATTA	ATAGTTCTTA	4020
TCTCAACACT	AAAGTACATA	AACATCAAAT	ATCAAAAATA	TATAAGAACA	ACATACTACA	4080
TTGTTTTAAT	GAAAACCTTA	AAAGGAATGG	TTAAACTCTC	ATTAAGCTAA	AACCAATGCA	4140
AAAATATCTT	TATAAATTAG	CAAAGAAGT	AAAAGTCACA	AACAACCTACC	ATAAAAATTT	4200
GGTAGTAAAT	TCTGGAAGT	AAATTTACTA	TAAACTCAAT	TATTCTAAAA	AAAATATTGC	4260
CTTAAATTAA	AGAATGCCTT	AAAAAAACAA	AATGCTCTGA	TTTAAACCTA	TACCCAAAAT	4320
ACAAATTTAC	TAAAGAAGAA	GATATAGATT	TAGAGAAGAT	CTTAATAATA	AAAATATTAA	4380
TATAAAAGTT	GCTCAGTATG	CTAAAGGCAA	AGAGTTTAAG	TCAAGTTTAG	AAATTACAAA	4440
GAGTAAAACT	ATAAACTTCC	TTTAAGAATG	AAAATTTATT	TTTATACTTA	CTTGGCTTAA	4500

TATTAAGATT	TTTTTATTCT	TTTCATAATA	ATCTCTTCTA	TCACTTAACA	TTTTGCTATA	4560
CAAAAATCTT	ACACATCTAA	ATACTTTTTA	AAAAAATTTG	ATTAGTGTTA	GAATATATTC	4620
TATATTTATA	AACTTTATTA	GCACTCATAA	TTTTACTAAA	TTAATATATT	ATATTTAATT	4680
TATTTTTTAA	ATTTATCTCC	ATTTACCAAA	AAAACATAAA	TAAAACCTCT	CAAACCTATA	4740
AATAAAAAAA	TAAGGCACAA	CCCCAACAAA	CTCAAGATCT	ATAATACAAA	AATACAATAT	4800
AAGAATCCCA	AGCTTAAAAA	CAACCCCTTA	AAATCTTTTT	TTATTGGCGT	TTTTAAATAA	4860
TGGTAATAAA	GAATTCCAAT	CAACACGATC	CCCCCTACAA	CTTTTCAAAC	CCTATAGCTT	4920
GGCTTTTTTAT	ATTATTTTTA	AATTTACATG	TCACAACAAT	AGATAATGCA	TAAAATAAGT	4980
ATTAATAAAA	CAAATACATT	TATAGAACCT	ATACAATTAT	TGAGCATATG	GCTAGTACTA	5040
AAAATGAAAA	TGTACAAGAT	AATATGCTAT	TAATAAAAAT	TAATGGCTAC	TAAAACTTTT	5100
GAATCCACAT	TTTTTCTTTA	AAAAAATCTT	AAATTATTAA	AATAAATAGA	AATTAATAAT	5160
ACCAAAAATA	TTATTATAGT	AATAAATATG	TAAAGCTATT	TTTATTAAAA	CTGATAATAA	5220
AAATATAATA	GCTAAAATAA	CATAAATTAA	CTTTAAATTA	TATCAAAGAC	TTAGATTTAA	5280
AATATTTAAT	AAAAGGCAAA	GCTATAAACA	CCATATACTT	ATTTTATTAT	TTTTTTCATT	5340
TTATTTAAAT	TAATTTAAAT	AAGACTCAAT	CAAATAATCA	ATCAAACATA	TTGGGTGAAG	5400
AAAAAATAGG	GTATTCTTGG	TGAATCGTTT	TAAAAGGGGG	TATAGTAAGC	TAAAAAACTC	5460
TTATTAAAGA	GGATGTTTAT	AGACTTAAAA	GTCTAATTCA	ATATGAAAGA	GGCTTTTTTAA	5520
AGCTAAAAAT	GTTAAAGAAA	ATCAAATTAA	GCAACAAGAT	GGTTTGTGTT	CTATAAATAG	5580
TTTTAAAGAA	TATATACATT	TGCACATACC	CTTCATTATA	ACATCTACTA	ATTACACAAT	5640
AAAAATAAAA	ATGATTTATT	AAGAATTATT	AGTAACTTAT	AAAAACTTTA	TAAGTTACAT	5700
AGTCAAAAAT	ATAAAAAATT	AAAACAAAAA	ATTAACGATA	TGGAAAAATT	GTATTTTATA	5760
GAAATAGAAA	TATATTTGCA	TTAAACAACCT	ATGAATTTAT	AAAGATTCTA	GTAGGAGAGA	5820
AAATATGAAA	AAAAAAAATT	TATCAATTTA	CATGATAATG	CTAATAAGTT	TATTATCATG	5880
TAATACAAGT	GACCCCAATG	AATTAACCTG	TAAAAAATG	CAAGACAAGA	ACGTGAAAAT	5940
TTTAGGATTT	TTAGAGAAAA	TTCAAGCAGA	TAATAAAGAA	ATTGTTGAAA	AACATATAGA	6000
AAAAAAGAA	AAACAAATGG	TGCAGGCTGC	TTCTGTAGCA	CCTATTAATG	TAGAGAGTAA	6060
TTTCCCATAT	TATCTTCAAG	AAGAAATAGA	GATAAAAGAA	GAAGAGTTGG	TTCCAAATAC	6120
TGATGAAGAA	AAGAAGGCAG	AGAAGGCAAT	TAGCGATGGG	AGTCTTGAAT	TTGCTAAATT	6180
AGTTGATGAT	GAAAATAAAC	TTAAAAATGA	ATCTGCGCAA	TTAGAATCTA	GTTTTAATAA	6240

TGTTTATAAA	GAAATCTTAG	AACTTGCAGA	TTTAATACAA	GCAGAGGTGC	ATGTTGCAGG	6300
AAGGATAAAT	AGCTATATAA	AAAAAAGAAA	GACCACTAAA	GAAAAAGAAT	ATAAGAAGAG	6360
AGAAATTAAG	AATAAGATAG	AAAAACAGGC	TCTAATTAAG	TTGTTCAATC	AGTTATTAGA	6420
AAAAAGAGGC	GATATTGAAA	ATCTTCATAC	TCAATTAAAT	AGTGGACTTA	GCGAGAGAGC	6480
ATCTGCAAAA	TACTTTTTTG	AGAAAGCCAA	AGAACTTTA	AAAGCTGCTA	TTACTGAAAG	6540
ATTAAATAAC	AAACGTAAAA	ATCGGCCATG	GTGGGCAAGA	AGAACACATA	GTAATTTAGC	6600
AATACAGGCA	AAAAATGAGG	CAGAGGATGC	TTTAAACCAA	TTAAGTACTT	CTTCTTTTAG	6660
GATACTTGAA	GCAATGAAAA	TAAAGGAAGA	TGTAAACAG	CTTCTTGAAG	AAGTAAAATC	6720
TTTTCTAGAT	TCTTCAAAGA	GCAAAATCTT	TTCTAGTGGC	GATAGATTAT	ATGATTTTTT	6780
AGAGACGAGT	AAATAAAAAA	ATATATTTTA	AAGGCTAATA	ACTTAAAATC	AAAGTCTTCT	6840
GTAAAGGAA	GACTTTTTTA	TAATTTTATT	TAAATAACGA	AAAGCTTGAT	AGTTAAAAAA	6900
TCTTTTTTAT	TAAAAATATG	TTTACTAAAC	AGAGCTCAAA	AATGACTATA	TTTAGTATCT	6960
CTATAAAAGA	ATTTTTCAAT	ATTTTAAAAA	ATTTATAGAT	AAACATAATC	TAAACCATG	7020
CATTAATACA	AACCTAAAAC	ATACTTGGTC	ACTTGTAATA	GTAAATTGTA	TCTAACTTTT	7080
TTTATTTATT	GAATATACGT	AAAAATTCTT	TATAATTTCT	ATTTTAAAAC	GCTGCTATTT	7140
AGCAATACAA	TAAAAGGCAT	TACAGATTGC	AATCAAACAA	ACTAAAGTTT	AAATAAAATA	7200
TTACCCTCTG	TTCTAATCCT	ATCAAACAAG	GTAATAAATT	CTTTAAATTT	CTAAAAGCCT	7260
AAACTTTAAA	AGAACTTGTC	GAAAAATAATA	TTTCTCTTAA	AAAAGGTTCT	AATCTTTTAT	7320
TTATAAGAAC	TTTTATACTA	TTATAAAAAAT	GTATCTTGCC	TTGATATATT	TGTATTCTTT	7380
ATAAATCAAG	CCTTCTACTT	TTTTTAAGAA	TATTTCTATT	TTTTATAAAC	TAGTTTTCTA	7440
CAATAGAAAA	GAAATAACCC	AAAGCCCTAA	AACTTAAAT	AAATGTTAGC	TATAATAACT	7500
AAAATAGAGA	TAAAAAACTC	AATCATAAAT	AATGGTAAAA	CAAACCTAAA	CCACGTACCA	7560
TAACTCAATC	TGGATATCCC	CAATACAGCC	ATTATAACTC	CGCTGGTAGG	TGTTATCAAA	7620
TTAATAAGCC	CAGATGCAGT	CTGCATGGCA	ATAACAACTG	AAGCTCTTGG	AATTGACAAA	7680
AAATCGGCAA	GAGGAGCCAT	TATTGGCATA	GTGAGACTAG	CATGTCCTGA	TGAAGATGGA	7740
ACAACAAATC	CTATAAATAT	TTGAATAATT	TCATTCAATA	TGATAAAAAG	GGGTCTTGGA	7800
AGATTGTATA	AAAAATTAGT	AGCAGCATTT	AACATAGTAT	CTGTAATCAA	CCCATCATCA	7860
CATACTATCA	TAACACCTCT	AGCAAGTCCA	ATAACAAGAG	CAGCGGTTAG	CAGACTTTCA	7920
GAACCTTTCA	CAAACGCATC	CCACATTTCA	GTTTCACCTA	ATTTACAAAT	AAAAGCCGAT	7980
ATAATAGCAA	CTCCAAGATA	CAACATTGTC	ATTTCTTGCA	TCCACCAACC	AAGATTAACA	8040

ATGCTAAATA	TCAAAATCAA	TATCATAAAT	CCAAATAAAA	GTAAAACTAA	TTTATGAGCA	8100
AAAGTAAACT	CAAGAGCATT	CTGAGCATT	TCTCCGGTAG	AAAGTCCATC	TTTTTTAACA	8160
AAATATTGAT	AATGTTTCATC	TTTTTGAGAA	TACACAAGCG	ATTTTGAGGG	ATCCTTTTTTA	8220
ATTTTAGACG	CATAAACACA	AACATAGGTT	ATAGCAGCCA	ATACTGATAC	AAAATAAAGA	8280
ACAATTCTAA	AATAAAATCC	ATCCTGCAAG	CTAATAGAAG	CTATTGCAGA	TGCAATTCCCT	8340
GTCGCAAATG	GATTTACAGT	AGAAGCCATA	GTTCCCACTC	CAGCTCCTAA	AGCAATAATA	8400
GCCGCTCCAA	CAAGACTATC	ATAACCCAAA	GCTACTATCA	AGGGAATCAT	AACAAAATAA	8460
AAAGGAAGGG	TCTCTTCACT	CATTCCGGTT	ACAGTTCCAC	CAATTGAAAA	AATAAACATT	8520
AACAAAAGGAA	TAAGCAACTT	ATCTTTGTGC	CCCAACTTCT	TGATTAAAAA	ATAAATTCCT	8580
ACATCTATTG	CTCCAGTTTT	CATAATAATC	CCATAAGCAC	CCCCAACAA	TAAAACAAAA	8640
ACAATAACTT	CAACTGCATG	TTCCATCCCC	TTTGACATTG	CGGTAAAAAT	AGTCATAATA	8700
GGATGTAAAA	ATCCCCTAGA	GCCTCGATCT	ACATATTGAT	AAGTTCCAGC	AACAATTATT	8760
TCCCTTTTAG	ATCCATCACC	CATTGCTTA	AATTCTTTAT	CAAACCTACC	GGCAGGAATC	8820
ACATACGTTA	AAATGGTAAC	AAATACAATT	AAAGAAAATA	TTATTGTAAA	ACTACTTGGC	8880
ATTTTGATCA	TAACGTTTCT	CCTAAATAAT	TTCATAAATT	TAATTTTACA	TAAAAAATAC	8940
TGTTATCCCA	AAGTTGATAC	CATAATAGCT	TTAATGGTAT	GCACCCTATT	TTCAGCCACA	9000
TCAAAAACAA	CTGAATTTTT	ACTTTCAAAA	ATTTCTTCTG	TAACCTCAAT	TCCATCAAGT	9060
CCGTATTTAT	CAAAAATATC	CTTACCAATC	ACAGTGTTTA	AGTCATGAAA	AGCAGGCAAG	9120
CAATGCATAA	ATATTGCATC	ATCTTTTGCC	ATGCACATTA	TCTCTTTATT	AACCTGATAA	9180
GCCTTTAGAA	GATTTATTCT	ATCTTCCCAA	TTACTCTCCC	CCATAGATAC	CCACACGTCT	9240
GTATACACAA	CATCAGCACA	TTTAACAGCC	TCTTCTTTAG	AATCTGTAAT	TGTAATTTTA	9300
CCCCCACTCT	CTAGGGCTAA	AGACCTAGCC	TTAAGCGTCA	AATCGGGGTC	TGGAAAAAGC	9360
TCTTTGGGAG	CAAAAATTCT	AAAATCAAGC	CCCATAATAG	CACAGCCTTT	CAATAAAGAA	9420
TTAGCAACAT	TCCCCCTACC	ATCGCCACAA	AACACTATTT	TAATCCCTTT	CAAACCTCCC	9480
TTATGTTCTT	TTATTGTCAT	TAAATCGGCT	AGTATTTGGG	TTGGGTGAGA	AATATCTGTC	9540
AATCCATTGT	AAACAGGAAC	ATTAGAATAA	TTCGCCAAAC	ATTCAACAGT	CTGTTGAGAA	9600
AAGCCTCTAA	ATCCAATAGC	ATCATACATG	CGTCCCAAAA	CTCTAGCGGT	ATCTATCATA	9660
GACTCTTTTG	AGCCCATTTG	ATTACCCTTA	GATCCCAAAT	AAGTAATATT	TGCCCTTGA	9720
TCATAGGCTG	CGATCTCAAA	AGCACACCGG	GTCCTTGTTG	AATCTTTCTC	GAAAATTATA	9780

ACTATATTTT	TACCTTTAAG	TTTTTGCACT	TCAATTCCCTG	CATATTTTGA	CTTTTTTAAA	9840
TTAATCGATA	AATCAAGTAA	ATATTTAATA	TCTTTGCTTG	TAAAATCTAA	AAGATTTAAA	9900
AAGCTCCTAT	TTCGTAAATT	ATACATCAAC	CACCAACCTT	TACAATCAAG	TTTTTAAAAA	9960
CTCATTTAAC	TCATGCTTAA	ACATGCTTAA	ATATTAAATA	TCCTCTCTTA	CTAAAGACAT	10020
AGACATGCAT	CTTGGCCCCAC	CACGACCCCT	TGAAAGCTCG	CTAGACGGAA	TTCTGTGAAC	10080
TTTAATACCA	TTTTCTTCAA	ACAGCTTATT	AGTTACATGA	TTTCTAGAAT	AAGCAATTAC	10140
TTCTCCTGGA	GCTATCGCCA	AAACATTAGC	ACCATCATTC	CATTGTTCTC	TTGCACCATG	10200
TATTAAATCT	CCACCCGCAC	ATTTTATTAT	GTCAATTTTT	CTGCCTAAAT	AAAAGCTCAA	10260
AACATCTTTA	AGCTTGGCTT	TTTCTTTTTT	AATATTAATT	TTATTAGAAT	TTGAATTGTA	10320
AGTTAAAACA	TAAATTGAGA	AATACATATC	ATCACTTGTA	AAACTTGTA	AAACGCTATA	10380
ATCAATTTGG	GTAAAACTG	TGTCTAAGTG	CATATAGGCT	CTGTTTTTTG	GAATTTTAAA	10440
AGCCAAAATT	GTGCTAAATG	GAGCCTTATT	TTTAAAAAGA	CTAGCAGCTA	GTTTTTCTAC	10500
AGACCCCGCT	TCTGTTCTTT	CTGAGATTCC	AATAACCAA	AGATCTTTAT	TTAAAACAAA	10560
CTCATCCCCA	CCTTCCAAAG	AAGTTTCTTC	CCATCTATTA	AACCAAATTG	GAACATTTTC	10620
TTTGTAAGCG	GAATGATATT	TAAAAATATA	CTCTGCAAAT	ATTGTCTCTC	TACGTCTAAC	10680
CTTGGTATAC	ATTTTATTTA	TTGTAATTCC	ATTGCCAATA	CTGGCAAAG	GATCTCTGGT	10740
AAATAAAACA	TTGGGCATAG	GATCAATAAC	AAAAAGACTT	GAACCATTAA	CCCAATCATC	10800
AAGCGAAAAT	TCACAATCTT	TAAGCTCTTC	TCTTGCAACG	CCGGAAATCA	TTTTAGAAAC	10860
CATATTATCA	ACGGTTAAAT	TAGAAAAATA	ATCTTTTAAA	ATATTAATTA	CACCATCTGT	10920
TTTTATTTCT	GCTTCCAGAA	TAAATTGAGA	TATAAATTTA	TTTTTGAGCG	CTACAGAAGA	10980
AGCAAGAACT	TCACTAACAA	GATCCTCAAC	ATACTCAATT	TCAACTGAAT	TATCTTTTAA	11040
AATATTTACA	AAAACCTCAT	GCTCTTGTCT	TGCAACTTTA	AGATAAGGAA	TATCATCAAA	11100
TAAAAAATTT	TTCATAATCA	AGGGTGTCAA	ATTTTCTAAT	TCTTCTCCTG	GCCTATGAAG	11160
CAAAACTTTT	TTCAAACGAC	CTATTTCCGA	AAATATATTT	ATTGGATTTA	AATATTCTTC	11220
TTCCATCGAT	TTCCCCCTTT	ATGAAAATTG	TCATATATTA	AAATACTATA	GTTTATATTA	11280
AAAAACATCA	ACTATTTTTA	ATAATATTAA	AAATATAATA	TAAATATAAA	AAATTGAAAA	11340
AATAAAAGTT	CTAAAAAACT	TCAAATCAAA	AACATAAACA	AAAAATTATG	CTAAAATACT	11400
AATCATGAAG	AATATTAATA	GATTAATATT	ATTAATATTA	ACTACACACA	CTTTATTATT	11460
CTCTTGCGCC	TTAATTGCAG	ATAATAAGTC	AAAAAATTTA	AGCACATCAG	AAATCATATT	11520
AACACAAAAA	ACACTACTAG	AAAGCTCTTT	AATAAAAAAT	CCTTCTAATG	TAGAATATCG	11580

AATACCAATA	TCCAGTATCC	AAGAAATTTT	AAACAATAAC	AATGATTCTT	TTTTAATAAA	11640
AAAAACAGCA	GCAAAAATCA	AAATAAGCCC	TCAAAAACCT	GAAGAAATAA	AAAACATCT	11700
AAATGCTTAT	AAAAATTATC	TAAATAATGA	AACAGAATGG	ATAAAGTTTA	TAGATCAAAG	11760
TAGCGTCAAT	GGAAATTTAA	CAATTAAAT	TGATACTGCT	TTTGAAAAAA	AAACAAATTT	11820
TAATCATACA	AATTCAGATA	ATGAAAATTT	AACAGAACTA	ATAGAACTAC	AAATGCATCT	11880
GGAAAAAGAA	ATTTTAAACT	TAATTGAGCA	AACATTTTCAT	GATAAAAATT	TAGGATATAT	11940
ACAATTAAGT	CACATCAACT	CATTCTTTCC	TCAAGAAAAT	ATAAACTCAA	TAACAAAAGA	12000
AATAATAGAT	GGAAAAGAAT	ATATTGCACC	GCACATAATA	GCAAATCAAT	TATTAAAAAT	12060
AAAAGATAAA	AAATATTTTG	AACAATTTAT	GCACTTTTTTA	AAAGTTGAAA	ACAGCAAAAT	12120
AAAAACAATA	ATTGAAAAAC	AAAAAATTTT	AGATCTTCAC	AATGAACTGT	ATTATTCAAA	12180
ACAATCCCCG	CCCAGAAGAA	GAAAAAGGTC	AACTGCCGAT	TCCGATAATA	ACAATAAATA	12240
CGATATAATA	CCAAAAATAA	TAGACCCAAA	TACAGGCATT	GAAATAACTC	CTAAAAATTT	12300
AAGATCTATT	TTATCAAATG	GCGACATAAT	ACTAATAAAA	CCAAAAATAG	ATTGGACAGA	12360
ATTTTTTTAT	TTTGGCAAC	ATGTGGGAAT	ATTTGATGAA	GAAAAATATG	AAGCCACTAA	12420
AAAAATTGCA	TTCAATGGAA	TTGATAGCTT	TGATATAAAA	TCAATAATTA	CAAGCAATCA	12480
AATCAAATTC	GATACAGCAT	CTACTCAAGG	TTCAGGATAC	GAAAAGCTTT	CAACATACGT	12540
ACAATCAAGA	ATATTAAAAA	TATTCTCACC	AATAACAGAC	ATAAGAACAA	TTCAAAAAGC	12600
TATTAATTTT	GGAAGAAGTA	GATACATTGA	CAATAACTTT	GGATATATGG	TTCCATTAAAT	12660
ATCCTCTAAT	TTATGGACAG	ATTCATTCAA	TCTTGAAGAA	ATTCACAACA	AAACCTATTG	12720
CTCTTTAATG	GTTGATAGAA	TATATAAAAT	AGCAGGACTT	AATGTATCAA	GAAATTACGA	12780
AATTTTCGGGA	ATAATTACTC	CTGGAGAAAT	AAATGCAGCA	GCTTACAATT	TTTACATGTC	12840
TTATACGATT	GCAGGAATAC	TTCCAAGCGT	GCTTCCAAAA	AGGCTCATT	AACCAACATT	12900
AAAAGAAAAA	TTCATTTGGT	ACAATAAAGA	AATAGTAGAT	GCAATAGAAT	TAAAAAATC	12960
GAAAGAAAAA	ATTTTGGGA	GAGCTTGCAA	CATTACAAAT	CTCTGGTGCT	CAGGAAGTTA	13020
ATACTACCCA	TGAACAAATA	TTATGCCTGT	ATTTTGATTA	ACAGCAATGA	AAAAATTATT	13080
TTCAAATCCT	GGGAAGAATG	CAAAACCGCT	ATTAAAGGAA	AAAACAATAA	AATAAAAAGC	13140
TTCAAAACAA	TAGAACAAGC	TCAAAATTGG	CTATTTAATA	ATGAGAATAA	AATTCACCAT	13200
CACCCAAATG	GAATATATTT	TGATTCTGGA	ACGGGAAGAG	GAAAGGGCAT	AGAAATTAGA	13260
GTTGTAAACG	AAAAAAGAAT	TTCAATATTG	GATAAAATCT	TAGATAAATC	CTTGATTAAT	13320

GAATATGGAA	ATTATTATGT	CAAAAATTTT	CAAGGAATTA	GCAATAATTT	TGGGGAAGCTG	13380
CTTGCCCTAT	ATACAGCTCT	CAAAATAGCA	TTAAAAGAAA	ATATAATAAA	CATATTTGGG	13440
GACAGTAAAT	TAATAATTGA	CTATTGGTCA	AAAGGAATCT	ATAATAGCAA	AAAATTAACA	13500
CAAATTACTA	TTAATTTAAT	CAAAAAGACA	ACTGAAGTAA	GGAAAAAATT	TGAAGAACAA	13560
GGTGGAAAAA	TTTCTTTTAT	TCCAGGAAAT	GAAAATATTG	CAGATCTTGG	TTTTCATAAA	13620
ACTAAGTAGA	AATATTGTCA	AAAAATACAT	AAAAACAATA	TTTCTGATTT	CAATGGTTTA	13680
TTTTTATTGT	TGTACGACAA	TAAAAATAAA	CCATGATTAT	GAAACTGATT	TTAAAGTTCT	13740
AGAATCTCCC	TCTAAATACA	TCAATATAGA	TGTAATTAAA	GCTACAAATG	AATATATTTA	13800
TATTCAAATT	ACAAACAATA	GCTTAGACGT	AGTAAAAATA	AATTGGCAAA	ACACTAGTCT	13860
TAACAACGAT	AAGATCGTCT	TAAAAAAGA	AGATCTTACA	ATAACAATG	AAACAGGGTA	13920
TAAAAATAAA	TACAGAGAGT	TTTTTATTGG	TCCTAAACT	TCATTTAAAT	TTAAAGTATA	13980
TCCACTAAAA	ATTCATTCTA	AAAACAAAAA	TAGCAATAAC	TTAAGCTCAA	CTATTAAATA	14040
TCCGTCTATT	TTTAAGCTCA	ACATAACAAA	AGTAGGAATT	GAAGCAAAAA	AAACAATAAA	14100
TGTTTTAATA	ACAAGAACTA	CAAAAATTAA	TATTACTAAT	AAATGAAAAT	CATTATTATT	14160
TTTTTGTTTT	TCTATTAATA	ATAAACTCGT	AGTTTGTATC	TTTGTTGTTT	TCAAATGACG	14220
CTTTTATAGG	AGCAAAAAAG	TTCCAAGAAA	TATAAATTGT	AAAATTATTT	CTCCAAATAG	14280
GAGAATAAAT	CCCATCAGAT	CCCCTTAATA	AATCTGGTTC	TAAATTATAT	TCTCCAACAA	14340
ATTTCCAATC	ATAAAAAATTG	AATTTAAAGC	CTGATGAAAA	TTTTTTAATT	TTAAAAAGTG	14400
AATCTTTTCT	GTCTTGAGAA	TTAAAGAAAT	TGAAAGATTT	TGATAAATCA	ACAAAGAAAT	14460
TAACAGGTTT	TAGACCAATT	TGGTCCATAT	ACCTTTTAAA	ATATTTAAAA	GTCTTAGTAT	14520
TAATAGATAA	AGTAGAAAAG	TAAATTTCTA	AAAATTCTGT	ATATTTAAAC	TTCAAAGTCA	14580
ATGCAGATCG	AAGTTCATTA	TCCGTAAATT	TCTGCAAATT	TATTTTCCAA	CCAACATCTA	14640
CCCCCAAGGT	AAAAGAAAGC	TTATTGTCAA	AAAAAGTTAA	AACGTACAAT	TCCTTTTTGT	14700
AACTAGAATC	TAAAGAATAT	GGAACAAGTT	TGGTTGTAGT	ACCAATCTTG	GAAAAATCTC	14760
CTTTTAAAGG	ATCATAATTA	TATTCAAAGT	CGTCTTTCAT	AGCAAACAAA	AATTGAAAAAT	14820
CAAAAACATT	AAGCTTAAAA	GAAAGTTCAG	AACTCTATT	TATCAAAGGA	TCATAGGCGA	14880
CTAAAAAAT	AAATTTAAAA	TAATCCAAAT	ATCTCGGCTC	AATTTTATAA	TACAAAGCAG	14940
GAGACATTTT	TAAATTTTAA	TAAGGCGATG	ATGGTTTTTG	AGGCTCCAAA	GGACTTTGAA	15000
CAGCAGAAAT	TCCAGAGTTT	TTCATAGCAT	CTTCTTTAAA	CTTTTATAAA	TATTTAATTC	15060
CAATCCCAGC	TTCTTGTAGC	AAATAAGGAA	AATCTAAAGA	AAGTTTAAGC	TCAGAAGAAG	15120

CTTTAATATC	TTCAAACTA	TTTTTTAATT	CACCTGAAAG	CTCAGTAGTA	AAATAATCAT	15180
AATCATAAAT	TAAAGAGGCT	GTTAAACTTT	GATAAAAAGT	TTCCGGATCA	GATAAAAAAA	15240
TACTACTATT	CTTATTAACC	AAAGATTTTA	CATCAGAATC	ATATTTTTTA	TTAAATGAAT	15300
ATAAAGTAGC	CTTATTTTCA	AACTTTAAAG	TACTTCTAGA	AAATAAAGGA	TATCTAATAA	15360
AAGGAAGCAA	GTTTAAATTT	ATTTGGTTAA	TAATAGAGTG	CTCACTTTTT	TTATCTTTAT	15420
CTTCAACTTT	AAAATCTTTA	TTTAAAGGAC	TATACTCAAT	AGTATTAAGA	TATAATAAAT	15480
TTTCAAAAGT	AATTAAACGA	TTGTAAAAAT	CAGCATGAAT	TTTTATATCC	GTTTTATTTT	15540
TTATATCAAA	TAAATAATTT	TTTATTTTCAT	AATTAAAGTC	CTTTGGACTT	GTTATGCCAT	15600
AATTATCAAA	AAAAACATTA	TTTCTTAAAT	AAGGATTAAT	GCCAAACCTA	ATAAAAAAAG	15660
AATCGGATTG	ATCAATATTT	TTTAAAGTAA	TTGGTTCTGG	AGGAATATAT	AAATCTTTGG	15720
TTAATTCTGT	TGTTTTTTTA	GTATTTTCT	CCTTCACACT	CTTTTTATCA	TTATCTTTAT	15780
CTTCTAGATT	TTTAATTTCT	GGGCGCATT	TCATTTCTTT	AGTATCAGCT	GGAAATGTCC	15840
ATTGGTTATT	GTAAAGATCT	TTTTGAAAAT	TCAAATCAAT	ATATGGAGCA	TAAATTCTCT	15900
CCAAATAAAA	CCATTTTCTT	GTAGGATCAT	TAACATCTTT	TGGTTTCTCT	AAAGGAGATT	15960
TAACATAAAG	ATTTTCATAG	CCCGACAATT	TAAAACTTAA	ACCTAAATTA	TTTAATTTAT	16020
AATCTAAAAT	CGAACCGTCA	TTAAATGTTT	GCTTATAAAA	AGAAGATAAA	TTCCAATCAA	16080
AAGTGCTAAT	GCTAGTTTGC	TCTTTAACCG	AATCTTTATC	TAAATTTAAA	AGAGAAAAAA	16140
ATGTAGCACT	TTCTATCCTA	TCTCTAAAT	CAATATTAAC	ATACGGGTCA	GAATAGTGCT	16200
CTAAAACAAC	CGAGAAAAGT	GCATCACTTA	AAAGAAATTC	TGTTTTTAAAT	TTAAATAAAT	16260
ATCTAAAAGG	AACTTCAAAC	CAAATACAT	CTCCTTGTT	AAGATTGGAA	AACTAAAAAA	16320
GAGATTGTTT	TAAAGTCCTA	TTATCAAAAG	GATAATATCC	TCCATCGTAA	CTATAAACAT	16380
TCCTGGTAAA	ACCCAATCCA	AAATTTCTTT	CCAAAGTTTT	AAAATGCCCC	AAAGTATTGC	16440
CCAAATTAAA	ATCAATTCCA	GAATAAAATC	CCAGATTAGC	ATAAATGTCA	AAAATAAGCT	16500
TAACATAATC	TTTATTAACA	CTGGGTGCTA	AATTTTCTGC	AAAAAATAA	GTAAATATC	16560
CATTTCTTAT	ATAAGGTTTT	TTACCCGAAT	TATAAACAGA	ATTGAAATCA	AAATCCAAAA	16620
AAGAAGAATC	TTCACTTGAA	GATTTATTAC	CAAAAAGATA	AACGGTATTA	AAAACAGAAA	16680
AACCTTTTCG	TGGATTTAGA	CCTAAAGATG	GATTAAAAAA	CAAACTATCT	CCCGGTCTGA	16740
AAAAAAAAGG	AATATAAAAT	ACTGGAACCTC	TTCCCATGTA	AAATATGGCA	TTTAAAAACC	16800
CAAAATCTCC	CGAGGGCAAT	GCCCATATTT	TAGAAGCCTT	GATTGAATAG	TAAGGCTCTG	16860

GAATTTTACT	AGTTGTTGCA	AAAGCTTGTT	CCAAAATGGT	AACATCATTG	TCTATCTTTT	16920
TTAAAACCTT	TCCTCCAAAC	GAAAGAATAT	GATCTATTTG	ATTTTTTTGC	ATTTTTTTTT	16980
GAAGAATACC	ATTTTTTAAT	AAAAAATTTT	GAGAATCAAA	ATCGACAAGA	AATTCATTGC	17040
CATAAAAATA	AAGCTTTTCA	TTGGTATCCA	TATCAAGAAT	ATATTCAACA	TTTCCAATAG	17100
CATAAAGTTT	TTTAGAGTTC	TTATTAAGGA	CTATTCGTGC	GCCTTTAATA	TTGTGCTTTT	17160
TATTTTCTTT	AATATCTTCA	ACCAAGATAT	TAACTCTTCC	TTCAAAAATA	ATACTTTCAT	17220
CTTTAGTAAG	TCCATAAGTG	AAATTTTCAA	GATTATCTGC	AGTTTCAATG	ATTATTTTAT	17280
ATCTACCAGA	TCCGGCAAGT	CCCTTTCCTT	TGATAAAAAG	CTCAGGATCT	ATTCCAAACT	17340
TTTTTAAAG	CAATTCTCGT	ATTTTGTAAA	CATCTGTTTC	TTTTAAACCC	TCTTTTAAGG	17400
CCCATTTTTT	TAAATCCTCA	TCGGTTGAAA	GCTCAAGTTC	TCTTAAATAA	GATTTTTTGAC	17460
TTAAAGTTAG	CTTATCCCTT	TTTTTAGAAT	TTTCATCATC	TATAGTCTGG	GCAAAAATTG	17520
CATTAGAAAA	TGTTAAAAAA	ATTAAAAATA	CTATAAAAGA	TTTTTTAAAA	ACATTCCCTG	17580
ATAGGAATTC	TCGCATTTTG	CAACCTCTTC	AGGAATACCA	GAAACAACGA	TATTTCCCCC	17640
TGCCAACCCA	CCATCAGGAC	CCAAATCTAT	TATATAATCT	GCCTGTTTAA	TTACATCCAA	17700
ATTATGCTCT	ATTAGTACAA	CTGTATTACC	ATTGGAAACT	AACCGCTGCA	AAACCTCTAA	17760
CAACTTCTTT	ATGTCATCAA	AATGCAGCCC	AGTTGTTGGT	TCATCAATAA	TATAAAAGGT	17820
TTTACCCGTG	CTCTTTTTAC	TTAACTCAAA	AGCCAACCTA	ATGCGCTGAG	CTTCTCCTCC	17880
TGATAAAGTT	GTTGCAGATT	GTCCTAATTT	AATATATTCA	AGTCCAACCT	CAATTAAAAA	17940
TTTTAAATAA	TGACTAATTT	TTGGGACATT	CTCAAAAAAT	TTACTTGCCT	CAAAAACACT	18000
CATCTCTAAA	ACATCATGTA	TATTTTTTCC	TTTGTATCTA	ACTTCTAAAG	TTTCTTCATT	18060
GAATTTTTTA	CCCTTACATA	AATCACAAGG	AACAAAAACA	TCTGGTAAAA	AATGCATTTG	18120
AATATTAAGA	TACCCATCTC	CTTGACATTT	CTCACACCTT	CCACCTTTAA	CATTAAAAGA	18180
AAATCTGCCG	GCTTTAAAAC	CCCTTGACTT	TGCATCTGGA	AGCTTGGCAA	AAAGCTCCCT	18240
AATTTCTGTA	AAAAATCCAA	CATAAGTTGC	TGGGTTTGAT	CTTGAAGTTC	TCCCTATTGG	18300
TTTTTGATTT	ATTTGAATAA	TTTTATCGAT	TTTTTCATAC	CCAACAATAT	CTTTAAAGCC	18360
ATCACAATAC	TTTTCATTA	GCTTTAATCT	ACTATCAAGA	GCTGGATATA	ACACCTCGTT	18420
AAGTAAAGTA	CTTTTCCGCG	TACCAGAAAC	ACCTGTTATT	ACGGTAAAAA	CTCCCAAAGG	18480
GATACTTAAG	TCTATATTTT	TAAGATTATT	TTTATTAGAG	CCCAAAAGCA	AAATTTCTCC	18540
CTTATCTGCC	TTTCTTCTAG	AGCTTGGAAC	ATCTATTTTA	AACTTGCCGC	TAAGATATTG	18600
ACCAGTTAAA	CTATTTTTGC	TATTTAAAT	ATCAATCAAG	GCTCCCTTTG	CAACTATTTT	18660

CCCTCCAAGA	ATTCCAGCAC	CAGGACCCAT	ATCAATAATA	TAGTCCGCAG	TACGCAAAGT	18720
TTGCTCATCA	TGTTCAACAA	CAATTACAGT	ATTACCAAGA	TTTTTAAGAT	TAACAAGAGT	18780
AGAGATTAAT	TTTTCATTAT	CTCTTTGATG	AAGACCAATA	CTTGGCTCAT	CAAGAACATA	18840
AATAACACCC	GAAAGTGCTG	ATCCTATTTG	AGTAGCAAGC	CTAATACGCT	GAGCCTCGCC	18900
ACCAGATAGA	CTACCTGATA	TTCTATTTAA	ATATAAATAA	GAAAGGCCAA	CATCAATTAA	18960
AAATTTAAGC	CTACTTTTAA	TTTCCTTTAA	AATTTCTTTA	GATATTTTTT	CGTCCACCAT	19020
ATCAAGCTGC	AAGTTTTCOA	AAAATACATA	AGAATCAAAT	ACTGACAAAT	TGGTAAGATC	19080
TTGAATGTCT	TTTCCATTAA	TTTTTCACAGT	TAAAGCTCCA	ACGCTTAGGC	GTTTACCTTT	19140
GCATGAATTA	CATATTTTTT	TAGACATCAA	ATTTTCGTAA	AAAATTTTAG	TACTCTCTGA	19200
TTCTGTTGCA	AGATATCGCC	TTTTTAAAAG	GGGCAAAAGT	CCTTCAAATG	TTTTAGAATA	19260
ATGAAATCCT	CCATCTAGCT	CTTTTGCTTC	CATTTCTTTG	GACTGGTAAA	TAAAATCTAT	19320
TTTTTCATTT	GAGCCGTATA	AAATCTGTTT	AAGAACTTTA	TCTGGAATGT	CTTTTATGGG	19380
AGTATTTAAG	TCAAAATTAT	AATGTTTAGC	AAGTCCTTTA	AAAATAGCCA	CAGACCAAGA	19440
TGAACTTGTC	TTAAACGTAA	GAAAAGCATC	ATCATTAAAA	GAAAGACTAG	TATCAGGACA	19500
AATGCTCTCA	AAATCAAAC	CAAGTGTAAC	GCCAAGACCA	GAGCACTCAC	TGCAAGCACC	19560
AAATGGACTA	TTAAATGAAA	AAAGTCTGGG	TTCTATAAGA	GGAAGTGAAA	ATCCACACAA	19620
AGGACAACTG	TTGTGCTCTG	TAAATAGTTT	GTCTATTTTT	TCCAAATCAT	TATCAATTTT	19680
CACTCTTAAA	TATCCATTAG	AAACAGCAAG	AGAAGTCTCA	ATAGATTCCG	CAAGTCTAAC	19740
TCGAACATTA	TTACCAAGCT	TAATCCTATC	AACTATAATT	TCAATGGTAT	GTTTTTTATT	19800
TTTATGTAAA	TTTAAATTAA	GTGCATCTTC	TATTAAATAA	TCTTCAGAAT	TTATCCTAAC	19860
TCTATTAAAA	CCTTGATTTA	ATATTTTTTC	TAAAACCTTT	TTATGAGAGC	CTTTAGACCC	19920
CCTTACAATT	GGTGCAAAAA	GTATAACCTT	GGATCCTTCA	GAATAACTTA	AAATAGTATT	19980
AACTATTTTA	TCTAAAGATT	GCTCTTCTAT	TAATCTACCA	TCATTTGGAC	AGTATGCTTT	20040
ACCAATTTTT	GCAAATATTA	GTCTATAGTA	ATCATAAATC	TCAGTAATTG	TTCCAACAGT	20100
AGAGCGGGGA	TTATTGCTTA	TTGTTCTCTG	CTCAATAGCT	ATAGAAGGAG	AAAGTCCATC	20160
TATATAATCA	ACATTGGGTT	TTTTCATTA	ACCTAAAAAC	TGCCTTGCAT	AAGCTGAAAC	20220
AGATTCCATA	TACCTTCTTT	GCCCTTCTGC	AAAAATAGTA	TCAAAAGCCA	GAGAAGACTT	20280
GCCAGAGCCA	CTCTTGCCAG	ATATTACAAC	TAAACCATCT	TTTGGAATAT	CTACATCAAC	20340
ATTTTTTAAA	TTATGTTCTT	TTGCTCCTCT	GACAATAATT	TTTTTTTTCA	AACTTTTTTC	20400

CAAAAATTAC	ACCTCTCTTT	TTTATTACG	AGCTATACTA	ATTTTGCTAC	TAAGCTCTTT	20460
TATTTTATCT	CTTAAAACAA	TTGCGTCTTC	AAATCTTTCA	TCATTAACAG	CTTCTTCTAA	20520
GTCAAATTTA	AGCTTATCAA	TAAGCTTTTT	TTTAGACAAT	CTCTCACCCG	AAATAATTTT	20580
TTCAAAATCA	TAGCCAACAT	TTTTATTTTT	ATTATTAAGT	TCCTTTTCTA	AAATATTTTG	20640
AATCTTTTTA	ACAATTGTCT	TAGGAGTAAT	ATTATTTTTT	TTATTATAAT	CAATCTGAAT	20700
TTGACGTCTT	CTATTAGTCT	CCTCAATTGC	CTCCCGCATA	GCTAAACTAA	TTTTGTGCGTA	20760
ATACATTATT	ACAAGTCCAT	TAGAATTTCT	AGCAGCCCTA	CCAATTGTTT	GTATTAATGA	20820
AGTAGTAGAT	CTTAAAAATC	CCACCTTATC	AGCATCTAAT	ATTGCAACAA	GAGATACTTC	20880
TGGAATATCT	AAGCCCTCTC	TAAGCAGGTT	AATCCCAACA	ATAACATCGA	TTTCAGATTT	20940
TCTAAGCAAC	GAAATAACTT	CCACTCTCTC	AAGGGTATCA	AGCTCTGAAT	GTAAATATTT	21000
TGCCCTTACG	CCAAGATTTA	CCAAATATTC	AGTCAAATCC	TCAGACATTT	TTTTTGTCAG	21060
AGTAGTAATT	AAAACCCGCT	CTTTAAGAGC	CACTCTTTTT	TGAATTTTCG	TGTAAAGATC	21120
TTCCATTTGC	CCATCAGAGT	GCCTAGTAAT	AATTTTCAGGA	TCAACAAGAC	CTGTTGGACG	21180
AATTATTTGG	TCAACAACCA	CACTACTTTT	CTCATCTCTT	TCAACACCCG	GGGTTGCAGA	21240
TACAAACACA	ACCTGATTAA	TTAATGCTTC	AAATTCATCA	TATTTAAGAG	GTCTGTTTTT	21300
AAGCGCTGCA	GGAAGTCTAA	ACCCAAAGTT	AACAAGATTT	AATTTTCTAG	AATGATCTCC	21360
ATTATACATT	CCCCTAAATT	GAGGCAATGT	AACATGAGAT	TCATCTACAA	ATAATAAGTA	21420
ATCTTTTCGA	AAAAAATCAA	AAAGACAATA	AGGTCTTTCC	ATTGTACTTC	CACTCAAATA	21480
TTTAGAATAA	TTTTCAATGC	CCGAACAAAA	CCCTGTTTCT	CTAAGCATTT	CCAAATCATA	21540
CTCTACCCTC	TGTTTGAGTC	TCTCGGCTTC	TACAAGTTTG	CCATTGTCTT	TAAAATATTG	21600
ACATTGAAGA	CTTAAATCAT	GAGATATTTT	GGGTATCGCT	TCTAATACAT	TTTCATAAGG	21660
AATTACAAAA	TAAGATTTAG	CAAAAAGAGT	AAACTATTTT	GTAGCTCCTA	AATTTTTTTT	21720
AGAAAATGAA	CTAACTCTAT	ATATTTCAAC	AATTTTCATCA	AAATCTAAAC	AAATTCGATA	21780
AGCAAACCTC	CCATGTTTAC	TGCTAGGCCA	AATTTCAACA	ATATCTCCCT	TAATCGAAAA	21840
TTTATCTCTT	TCTAGATTCA	TTAAAGTTCT	CTCATAATAA	AGCTCTACAA	AAATATCTGA	21900
TATTTCTTTA	ATAGAAATCT	TTTGACCTAC	AAAAAATTCT	CGTGCTGATT	TTTTGAAAAA	21960
ATCTGGAGAT	CCAAGAGCAT	AAATTGAAGA	TACGGTTGCA	ACAACAATTA	CATCTCGTCT	22020
TTTAGCAAGA	GACGTTACCG	TTCTTATTCG	CTTAATTTCT	ATCTCAGTAT	TAATAGTGGC	22080
TTCTTTTTCA	ATAAATAAAT	CTTTTGAAGG	AACATAAGAT	TCTGGCTGAT	AATAATCATA	22140
ATAAGAAACA	AAATACTCAA	CAGCATTATT	TGGAAAAAAA	TCTTTAAACT	CTCTATAAAG	22200

CTGTGCTGCT	AATGTTTTGT	TGTGACTGAC	AACTAAGGCA	GGCCTGTTTA	GATCTTTTAT	22260
TATATTTGCA	ATTGTAAAAG	TCTTTCCACT	GCCTGTAACA	CCTTTTAAAG	TTTGATATTT	22320
ATTTCCAAGC	AAAATAGAAT	TTTCAATCTC	TTTTATTGCC	TTAGGCTGAT	CCCCAGCAGG	22380
AAGATATTCT	GACTTCAAAA	AAAAATCTAT	CATTAATTTA	ACGACCAAAA	TTTAATACAC	22440
ATTCTTATAA	ATTATATGAT	AATAAATTCT	ATATCAAGTA	TATAATTCAT	TATAAATCAA	22500
TATAATTTAA	TTAATCTTTG	TTTAATAAAA	TAAAAGGAAA	TATTGATGCT	AAAAATCGAA	22560
GCTAAAAGAA	AATTGAAAAA	TTATATTCTT	CTTGAAGAAG	ATATGCATTT	TAAAGAAGAA	22620
GCAATAAAAA	TTCAAAAAAC	AAATAATTCA	ACAGAAATTT	TAAATAGATT	TTACAAAGAT	22680
CTAGAATTTG	GCACTGCTGG	AATAAGGGGA	ATCATTGGAG	CTGGAACATG	TTACATGAAC	22740
ACATATAATA	TAAAAAAAAT	AAGCCAAGGA	ATATGCAATT	ACATACTTAA	AATAAACAAA	22800
AACCCTAAAG	TTGCAATAAG	CTATGATTCA	AGATATTTTT	CAAAAGAATT	TGCTTACAAT	22860
GCTGCTCAAA	TTTTTGCCTC	AAATAATTTT	GAAACATATA	TATATAAAAG	TTTAAGACCT	22920
TCCCCACAAC	TATCTTATAC	AATAAGAAAA	TTTGACTGTG	ATGCTGGCGT	TATGATAACA	22980
GCAAGTCATA	ATTCAAAAGA	ATATAATGGA	TATAAAGCAT	ATTGGAAAGG	TGGAATCCAA	23040
ATAATACCAC	CTCATGACAC	ACTAATAACT	AATGAAATTA	AAAATACAAA	AAACATAATA	23100
AATACAATTA	CCATAAAAGA	AGGCATTGAA	AAAGGGATCA	TCAAAGAACT	TGGCAATGAA	23160
ATAGACGAAG	AGTATGTGAA	AGCAATAAAC	AAAGAATTGC	CTGATTTTGA	AAAGAATAGC	23220
AAAGAAACAA	ACTTAAAAAT	AGCCTACACA	GCATTACATG	GCACCGGTGG	GACCATAATA	23280
AAAAAACTCT	TTGCAAAATAG	CAAAATACGG	CTTTTTTTTAG	AAAAAAATCA	AATACTACCA	23340
AACCCTGAAT	TTCCAACAAT	AAATTATCCT	AATCCAGAAA	AACAAACATC	AATGCTTAAA	23400
GTAATAGAGC	TTGCAAAAAA	AAAAGATTGT	GACATTGCCC	TTGCAACAGA	TCCAGATGCC	23460
GACAGAATAG	GGATTGCATT	TAAAGATCAA	AACGAATGGA	TATTCTTAAA	CGGAAATCAA	23520
ATATCATGCA	TTTTAATGAA	CTATATACTC	TCAAAAGAAA	AAAATCCTAA	AAATACATTT	23580
GTAATATCAT	CGTTTGTAAC	AACACCAATG	CTAGAAAAAA	TTGCAAAAAA	ATATGGTTCT	23640
CAAATTTTTA	GAACCTACAC	AGGATTTAAA	TGGATAGGAA	GCTTAATTAA	TGAAATGGAA	23700
AAAAATGAAC	CAAATAAAAA	ATTTGCTTTT	GCATGCGAAG	AAAGTCATGG	ATATCTAATA	23760
GGAAGAAAGG	TTAGAGATAA	GGATGCATTT	TCAGCCATAA	AAGGAATTTG	TTCTTTAGCA	23820
CTTGACTTAA	AAGCCAAACA	ACAAACAATT	AAGGATTATC	TTGAAAAGAT	ATACAAAGAA	23880
TTTGGATATT	ATGAAGAATT	TAATATAGAA	AAAAACTTTG	AGGGGGCCAA	TGGAGAAATT	23940

CAAAGAGAAA	AGTTAATGCT	AAAACATAAGA	AAAGAACAAA	AAGTACAATT	TGCAGGAATT	24000
AAAATAATTG	AAAAATTAGA	CTATAAACT	CTTAAAAAGA	TTAACTTTAA	AAATGAAATT	24060
TCAGAAATTA	AAGAATATAA	ATACCCCAT	AACGCAATAA	AATTTATACT	TGAAAACGAA	24120
ATTGCAATAA	TTGTAAGACC	CTCTGGAACA	GAGCCGAAAA	TTAAATTTTA	CATATCTGTA	24180
AAACTAGAGT	ATAAGGAAAA	ACATAAAATA	TTTGATATAA	TAAATGCAAT	AAAGATGGAG	24240
ATAAAAAAAT	ATTAACATAA	CAGAAAATTT	AATAAATTTG	GTAGAAATAG	ACTCAAAAGA	24300
AATTGCAAGA	AAAAATAAAA	ATAAGAGGT	TTCAATTTGG	CACTTATTAA	TGTCTATAAT	24360
TACCACTCCC	AAAAAATCCG	AAATAAAATT	TATAGATAGC	AAAACCTCTA	AAAACATTAA	24420
ACAAGAAGTT	ATATCTGAAA	TAGATAAATT	AGAGAAAATT	TTAATAGAAA	AAAACGAAAT	24480
AATTATTCCC	AAAATCAATA	AAGAAATCTT	TGCTCTCATA	AAAGAAGCTA	AAAAGGAATT	24540
TAAATCCAAA	CCTTTAATAG	GGGCAAAAGA	AATTTTTTAT	CAAATATTAA	AAAATAAAAA	24600
ACTTCTTAAA	AAACATAAAC	TAAGTAAATC	TAGCTTTAAC	TTTAAAGATC	AAAATATATT	24660
AGAATACATG	GAAAAAAATA	AAATAAGATT	AATTGAAACC	TACAAAGAAT	TTGATGAAGA	24720
AATACGACTT	GAAAATGAGC	ACTTTGAAAT	TGGAAAGTAT	GTCAAAAATT	TAACAGCACT	24780
TGCAAAAGCC	AAAAAATTAG	ACCCCTTGGT	TGGAAGAGAA	GCAGAGATTA	AAACTCTTAC	24840
AAATATACTC	TTGAGAAGAA	ATAAAAATAG	TGCAATGCTA	ATAGGCGAAC	CTGGTGTTGG	24900
AAAAACAGCA	ATAGTTGAAG	GCCTTGATC	AAGCATAGTG	CAAAAAAAAA	TAAGTAGCAA	24960
ACTACAAGAC	AAAACAATTC	TAATGCTTAA	GGTTTCAAAC	TTGGTATCGG	GAACAAAATA	25020
TAGAGGCGAG	TTTGAAGATC	GTTTAAATAA	TATAATTAAG	TATATTGAAA	AAAACAAAAA	25080
CACAATCATA	TTTATTGACG	AAATACACAC	TCTAATAGGA	GCTGGAAACT	CTGAAGGAGC	25140
TCTTGATGCA	TCAAATATAC	TAAAACCATC	ACTTTCTAGA	GCTGAAATAC	AAATTATTGG	25200
CGCAACTACT	TACAATGAAT	ATCGAAAATA	TATTTCAAAA	GACAAAGCAT	TCGCCAGAAG	25260
ATTCCAAACA	ATTACCGTAA	AAGAGCCTGA	TGAAAAGAT	aCACTAAAAA	TAATCGAAAA	25320
TATTGCAAAA	AATTTTGAAG	ACTATCATGG	AGTGATCTAT	GAAAAAAGCG	CGCTTTTAAA	25380
TATAGTAAAA	CTTTCATCCA	AATATCTAAT	AAATAAAAGA	TTTCCAGATA	AAGCAATAGA	25440
TATAATAGAC	ATTGCCGGCG	CAATTAAAAA	GGAAGAACTT	ACAAAAGACA	ACATCATAAC	25500
ATCAGATGAT	ATACAAAAGG	CAATAAATGA	AATATTATCT	ATTAAAACAG	CAAATAACAC	25560
TAAAGAAGAA	ATTTTAGAAT	TAAAAGAAAT	AGAAAGCGAA	ATAAATAAAA	AGGTGATCGG	25620
ACAAAAACAT	GCGGTAAGCG	AACTTATCAA	AGAAATTATT	AAAGTCAAAC	TTGGACTTAA	25680
TGACGATTCT	AAGCCTTTAA	CTTCAATATT	GTTAATAGGA	TCAAGTGGAT	GTGGAAAAAC	25740

TGCTTTAACT	GATGAAATAT	CTAAAAAAT	TATCAAAGAT	CAAATTCAG	TATTAAACT	25800
AGATATGTCA	GACTATAAAG	AAGAAAAC	TATTTCAAAA	TTAATTGGCA	CAAATCCAGG	25860
ATACGTAGGC	TACTCTGATG	GAGGCATTCT	GACAAATAAA	TTAAGACATT	CATTTGAAAC	25920
TTTAATATTG	TTTGAAAATA	TTGAAAATGC	CCACAGCTCT	GTATTAAACC	TAATAAGTCG	25980
AATGCTTGAA	AACGGAGAAC	TTATTGACAG	CAAAGAAGAT	AAAATACTAT	TTAAAAACAC	26040
AATTATAATA	ATGACTACAA	ACATTGGATC	TAGAATGCTT	CTTGGAGAAA	AAAATATTGG	26100
ATTCAACAAA	AATCAACAAA	AAAGCTTAGA	AACAAAAAGC	TTTAAAGAAG	AAATAAACCA	26160
AGATCTTGAA	AAAAGATTTA	AATTATCCTT	TTTAGACAGA	ATTCAAAAAA	AAATCATCCT	26220
AAATATCCTT	ACaAAGGAAA	ATGTAGAAGA	AATTTGCAAA	AACTACTTAA	ACACCCTTAA	26280
AACAAAATTT	CACTCTAAAG	GAATCGAGAT	AGAAATAAAA	AAAGATGTTG	ACAAATTCAT	26340
AACCACAAAA	TACTATAAAA	AAAAATCAGG	AGCAAGAAGC	GTAATTGCTG	CAATAAAGGG	26400
GAAAATAGAA	GAAAATATTA	TCACCAAAAT	AGCTGAAAAT	CAAAACATAA	ATAAAATAAC	26460
GATTTATTTA	GAAAAAGAAA	AAATAATAAT	AGAATAAAGA	GGAATTATAA	TATGTTTAAA	26520
AAAGTAGAAA	ACAAGGCAAA	TTTTCCTAAA	ATAGAAGAAA	AAATATTAAA	ATTTTGGAAT	26580
GACAATAAGA	TCTTTGAAAA	ATCAATAAAG	CAGAGAGAAG	GATGTGAAGA	ATTTACATTT	26640
TATGACGGAC	CGCCTTTTGC	AACAGGACTT	CCTCATTTTG	GACATTTTGT	TCCAAACACA	26700
ATAAAAGACA	TAATTCCAAG	ATATCAAACA	ATGCAAGGCA	AGTATGTTAA	AAGAAATTTT	26760
GGATGGGATA	CTCACGGACT	ACCTGTTGAA	TACGAAGTAG	AAAAAAAATT	GGGAATTTCT	26820
GGAAAATACG	AAATAGAAAA	TTATGGCATT	GAAAATTTTA	ACAAAGAATG	CAGAAAAATA	26880
GTACTTAGAT	ATACAGAAGA	ATGGAAAAAT	ATAATCTTGA	GACTTGACG	ATGGGTAGAT	26940
TTTGAAAAGG	GTTACAAAAC	CATGGATATA	AGCTTCATGG	AATCCGTGTG	GTGGGTATTT	27000
AAAAATCTTT	ATGAAAAAGG	TTAATCTAC	GAAAGTTACT	ATGTACTACC	CTATCCCCA	27060
AAGCTTGCAA	CTCCGCTTTC	AAATTTCGAA	GTGAATCTTG	GAGAATATAA	AGAAGTCAAT	27120
GACCCATCAT	TAACAATAAA	ATTTAAAATA	AAAGATAAAA	ACGAATACTT	ACTAGTGTGG	27180
ACAACCACCC	CCTGGACATT	GCCCTCAAAC	CTTGGAATTG	CAGTAGGACA	AGAAATAGAA	27240
TATTCTAAAA	TTTTTGACAA	AACAAAAGAA	GAGATTTTAA	TACTTGATC	AAAAAAGCTT	27300
AATAGCTATT	ACGATGATGA	AAATTCATAT	ACTATTATAG	AAAAATTCAA	AGGCAGCAAG	27360
CTTGAAGGCA	TAGAATATGA	ACCTATTTTT	AACTACTTTT	TAGAACAAAA	AGATAAGGGG	27420
GCTTTCAAGG	TACACACAGC	TGATTATGTT	ACAACTGACG	ATGGAACAGG	AATTGTTTCAT	27480

ATTGCTCCTT	TTGGAGAAGA	AGACTACAGA	ATACTCAAAA	AACACACAAA	TGTCGATATA	27540
ATAGACCCCT	TAGATGCTGA	ATGTAAATTT	ACAAATCAGG	TAAAAGATTT	TAAAGGACTT	27600
TTTGTAAGAA	ATGCTGATAA	AAAAATAATA	GAAAACCTAA	AATTACGCAA	TTTTTTATTC	27660
AAAAGAGAAA	ATTATCTACA	CAGGTATCCA	TTTTGTTATA	GAACAAACTG	CCCAATTATT	27720
TACAGACCAA	TAAGTTCGTG	GTTTGTAAT	GTAGAAAAAA	TAAAAACCAA	ACTTTTAGAG	27780
GTAAATGAAA	AAATTAATTG	GATGCCAGCC	CATTTAAAAA	AAGGAAGATT	TGGAAAATGG	27840
TTAGAAAATG	CAAAAGATTG	GGCAATAAGC	AGAAACAGAT	TTTGGGGAAA	TCCAATTCCA	27900
ATTTGGATAT	GCTCAAAAAC	AGGAAAAAAA	ATTTGCATTG	GATCAAAAAA	AGAGCTTGAA	27960
AACCTATCTG	GCCAAAAAAT	CGAAGACTTA	CATAAGACC	AAATAGATAA	AATAACCTGG	28020
CCAAGCAAAG	ACGGTGGCAA	ATTTATCAGA	ACAAGCGAGG	TTCTCGATTG	TTGGTTTGAA	28080
TCTGGAGCAA	TGCCTTACGC	AAGCAACCAT	TATCCATTCA	CAAATGAAAT	TAATTTTAAA	28140
AATATATTTT	CTGCTGACTT	TATTGCAGAA	GGTCTAGATC	AAACAAGAGG	ATGGTTTAT	28200
ACTCTTACAA	TCCTGGGAAC	TGCTCTTTTT	GAAAACACAG	CATTCAAAAA	CGTTATTGTA	28260
AATGGACTTG	TGCTTTCAAG	CGATGGAAGA	AAAATGTCAA	AATCCTTTAA	AAATTATACA	28320
GACCCAATGC	AAGTAATAAA	CACCTTCGGA	GCTGATGCTT	TAAGGCTTTA	TTTAATAATG	28380
AGCCCTGTAG	TTAAAGCTGA	TGATTTAAAA	TATAGCGACA	ATGGAGTAAG	AGACGTTCTT	28440
AAAAATATAA	TAATACCCAT	TTGGAACGCT	TATTCATTTT	TCACAACTTA	TGCAATAATT	28500
GATAAATTCA	AACCTCCAAA	AAATCTCAGC	CTGGCTAAAA	ACAATAACCT	TGACAAATGG	28560
ATCATAAGCG	AACTTGAAAG	TCTAAAAAAA	ATACTAAATA	CAGAAATAGA	CAAATACAAT	28620
CTAACAAAAT	CAATAGAATC	TTTACTTGAA	TTTATAGATA	AATTAAACAA	TTGGTACATA	28680
AGAAGATCAA	GGCGAAGATT	TTGGAAATCA	GAAAACGATA	AAGACAAAAA	TGATGCCTAC	28740
GAAACATTAT	ATTATGCAAT	CAAACTTTA	ATGATTTTAC	TTGCACCTTT	TATTCCATTT	28800
ATAACAGAAG	AGATTTATCA	AAATTTAAAA	ACTGATGAAG	ACAAACAATC	AATACACCTT	28860
AACGATTATC	CAAAAGCAAA	TGAAAATTTT	ATTAACAAAA	CAATTGAAGA	GAAAATAAAT	28920
CTCGCAAGAA	AAATAACTTC	AATGGCAAGA	TCACTCAGAT	CATTGCACAA	TATAAAAATA	28980
CGCATGCCTA	TTAGTACGAT	ATATATCGTC	ACAAAAAATC	AAAATGAACA	AAATATGCTA	29040
ATGGAAATGC	AAGAAATAAT	ATTAGATGAA	ATAAATGCAA	AAGAAATGAA	AATAAAAGCT	29100
AACGAAGAGG	AGCTTATAAC	TTACAAAGCA	AAAGCAAAC	TTAAAGAAGT	TGGGAAAAAG	29160
CTTGGAAGAG	ATATGAAAGC	GGTATCTACT	GAAATTAGCA	AGCTAAAAAA	TGAAGACATA	29220
ATAAAAATAA	TAAATGGAAC	ATCCTACGAG	ATAAAAGTAG	CCAATGCAAA	GCATTATTTA	29280

TCATTAAATG	ATATAATATT	AGAAAGAGAA	GAAAAAGAGA	ACTTAAAAGT	AATAAATGAA	29340
GAATCCATTA	CAATAGGAAT	AGACTCACTA	ATCACTAAAG	AGTTGTACTT	GGAAGGGCTG	29400
ACAAGAGAAT	TTGTAAGGCA	AATACAAAAT	TTAAGAAAAG	AAAAAAATTT	TGATGTTAGC	29460
GATAGAATAA	ATTTATACAT	AGAAAATAAT	GAAACTTTGA	AAGAAATGCT	AAATAAATTT	29520
GAAAAATACA	TTAAACTGA	AACATTAGCC	TTAAATATCA	TATTAAACAA	AAGTAAGCTA	29580
GAAAAAAAAA	TAAACCTTGC	CGATGACATA	TTTACACTAA	TAGGAATTGA	AAAATGTTAA	29640
AAACATTAAC	AAAAATAATT	ACCATTTTCAT	GCCTCATAGT	GGGATGCGCA	AGCCTGCCTT	29700
ACACTCCTCC	AAAACAAAAT	CTAAATTACT	TAATGGAAct	TTTACCTGGC	GCAAATTTAT	29760
ACGCCCATGT	AAATTTAATT	AAAAACAGGT	CTATTTATAA	CTCTTTAAGC	CCTAAATATA	29820
AATCAGTTCT	TGGGCTTATA	AGCAATTTAT	ACTTTAGCTA	TAAAAAAGAA	AATAACGATT	29880
TTGCTCTACT	AATAATGGGT	AATTTCCCAA	AAGATATTTT	CTGGGGAATT	CATAAAAATA	29940
GAAATACAGA	ATCAATAGGC	AATATATTTA	CAAATCCAAA	ATGGAAACTT	AAAAATTCAA	30000
ATATATACAT	TATTCCAAAC	AAAGCTAGAA	CTAGCATTGC	AATAACCCAA	AAAGATATAA	30060
CCGCAAAAGA	CAATAATATG	CTAACAACAA	AATATATTGG	GGAAATAGAA	AAAAATGAAA	30120
TGTTTTTTTTG	GATTCAAGAT	CCAACATTAT	TGCTCCCAAA	CCAAATAGTA	AGCAGCAAAA	30180
ATTTAATTCC	CTTTAGCAGT	GGAACCTTGT	CTATAAACAG	CTTAAATCAA	GAAGAATATA	30240
TTTTTAAATC	CTTAATCAAA	ACAAATAATC	CACCAATACT	AAAAATATTG	TCAAAAAAGT	30300
TAATTCCAAC	CGTCTTGACA	AACATGACAA	ACCTCACAAT	ATCAAGCCAC	ATAAAGACCA	30360
CAATAAAAGA	CCAAAATACG	GTTGAAATAG	AATTTAATAT	TCAAAAATCT	AGTGTGAAA	30420
GCCTTATAGA	AAAAC TAGCT	TCAAATATTC	AAACCTAAAA	TTTCTGCCAC	TCCACTAAAA	30480
TGAGGTATTA	TTTTGATTTT	TGCAAGTAAA	ATAATGAAAC	AAAGCTCCAA	TTTACCAGA	30540
TTCATTATTA	AATTTAGTAG	GCTCAAGTGC	TACAAGATTT	TTTATATTAT	TATTATTATC	30600
AAAAGCCATT	TCTAAAGACC	ATAAATTTTC	TAATTTTTCA	TATATTCTAT	CTATTAAATC	30660
GGGTCTTGCG	CTTATTCCTC	CTCCGATCAA	AATTTTTTCA	GGATTCAAAA	TAAAAGTTAA	30720
ATTAAAAATA	CCAAATGACA	AATTC TCAA	AAATCTATCA	ACTTCATTTT	TGGCATGAAT	30780
ATTCCCATT C	TCAGCTAGAT	CAAAAACAAA	TTCTCCTGAA	ACCTCTTTTA	AAGGTTTTCC	30840
TAATCGCATA	GCAACTCTTT	TTCTTAAAGC	CGAAACAGAC	GCAATGGATT	CCCATT TGCA	30900
ATTAAAGGGA	ATATTGTTGC	TAATACCTCC	AGTAATCATA	AATCCAACCT	CTCCGGACAT	30960
AAAAGAATTT	CCTCTTAAAA	GCTTGCCATT	TGCAAAAATT	CCAGCACCAA	TTCTGTGCC	31020

AAGAGTTATA	GCAATAAAAT	TATTAGAGTC	AATAGCATTA	CCCTTAAATT	TTTCTGCTAA	31080
GGCTACACAA	TTAGCATCAT	TTTCAATCTC	TGTACTTACT	CCGGTTAAAG	ATTCTAATCG	31140
CTCTTTTAAA	GGATAATTAA	CAAATCCAGA	AATAGCATTT	ACCCTAAGAA	CATTTCCCTT	31200
AAGATCAACA	AACCCAGGAA	TACAAATTGC	AACTCCCGCA	ATATCACTTG	ATTCTTTGTA	31260
AGAATTAATA	ATATTAACTA	AAATATTTAC	TTGTTCGTCA	GAAGTAGCAC	CTGTGCTTAT	31320
TTCAATTTTA	TCAAAAAAAA	CACCGCTTGA	ATCTGAAAGC	GAATATTTGG	TACTAGTCCC	31380
GCCAATATCA	ATCGCTAAAT	AATGTTTCAT	ATTTATCCTC	AAGGCCAAT	TGTACATAAA	31440
TGATCATAAA	TTTTCTATAG	CAATATAAGA	AATTTTAGAA	ATCTTGTTTA	TAACATTTTG	31500
CGCTTTAATC	ATCTCCTTCA	AATAAGAAAC	ATGGGAAATT	ATGCCAATTT	GTCGCCCAGT	31560
CATCATTTGA	AACTTAGAAA	GCTTAGGCAT	AACTTGAGCC	AAAGTATCTT	CATCAAGATT	31620
GCCAAAACCT	TCATCTAGAA	AAAAAGCCTC	TATTTTTAAC	TCATATCCC	TTATTTTATC	31680
AGATAAAGCT	AAGGACAAAG	CTAAAGACAC	AAGAAATTC	TCACCCCCAG	ACAAAGTTTT	31740
TACCGTTCTT	ATTTTATTAA	CATCTTTTTT	GTCTTCAATT	AAAAAATCAA	ACTCTTTGCT	31800
CTCTTTGTTC	GTTTTGAGCT	CAAAATCAGG	AAAAATCCAC	CTTAAATACT	TTTCATTTGC	31860
CAGTCTTAAA	ATATCATTA	TTAAAAAAGT	TTGAACATA	TATTTCAATC	CAGAAGATCT	31920
AATAACGACC	TTCTTAATA	CATCTAGCTT	ATCTTTCCTT	TCTTTAGCAA	GATTTAATTC	31980
ACCCCTTAGC	GAGTCTAAAT	TAATTTTTTG	TTGATTAATC	TTTTTTTGAA	GAGTTTGAAA	32040
ATTTAAAAGC	TTGATTTTAT	ACTTTTCAAT	ATCTCTGGAT	AAAAATTCAA	GCTTAGAACT	32100
AATTGATAAA	CTCAATTTTT	GCAAAAAAGA	CAGACTATTC	TTATTAATTT	GCTCTAAGTC	32160
AGTTGTTGAT	TCAAAAGAAA	ATGAACATA	AAAAACATTT	TTTAAATTTG	AAATTAGATT	32220
AATAAAATTA	TTTTGCTCTT	CATTTAATCT	CGCTTTTAAA	GTCAATATAG	ATTCCTTTGT	32280
AAATTTAATT	TGTGTTTCGG	TTTTAATTTT	TAAATCTTCT	AAATTTTTTAA	GATTTAAAAC	32340
AAGCATATTC	CATTCATTTT	CCACACTTTT	CTGCTTAGCC	AAAACAACAT	TAAATCCCTT	32400
TTCCAAAGAA	GAATAATCAT	TAAAACTTAA	ATTTAAATTA	AGTTTAAACA	ATAAATCCTT	32460
AATCTTTAAA	AGATTTTGAT	CAAAATTTTT	ATTTTTCAAA	GAAATTTCAA	TTTTTAAATC	32520
TTTTTCTTTA	GCCTTAAATT	GTTCAAGCTT	TTCTAATTTA	TTTTCAAATG	CTAAAATTTT	32580
TTCTCTGTCA	AAATAATTTA	TGTATTTGTC	AAATAAATTT	TTTCCAATCA	ATCTTAAAAT	32640
TTCAGCATTA	TTCTTTTCAA	ACTCCAAAGC	ATTAACCTCT	TGTTTAGAGA	TTTCTTCTTG	32700
TCTACAAGAT	AGCTGATATT	TAAGCTCATC	AATTTGATTT	TGTATCAAAT	GAAGCTTTGA	32760
ATTTAAAGCG	TAAAGCTCTT	TCAAACTGCA	AAGCGATAAT	TTATCTTTAT	GCTGATAATT	32820

TTTATATTCA	GCCTCAATAT	ATTTTAGCTT	CTCTTTATCA	CTCTCAATAG	AAAAATTTT	32880
ATCATCAAGA	TATTTTAACA	ACTCTTTATA	CAAGTCAATC	TTAATAATAT	TTTTTTCCTT	32940
ATTTTTATTA	AAATCTTCTT	TGCCAGTAGA	TTTAAATAAA	AATTCTAGCC	TATCTCTATA	33000
TTTTGAAATC	AACTCATCAT	TAAAAGTTTG	GAACAATTTT	AATGCTTCAT	AGTAAACATA	33060
TTTGTCAAAA	TCAAAATTAG	ATTTCTCAGA	AGAAATGCTT	TTAATCTCAT	CATTTTTTTT	33120
GCTCTGAGAT	CTTAATAATT	CTTTTTTTTC	ATCCTCAAGA	CTATTTTTCC	TTAAAAGCAA	33180
GCTTTGATAA	TCATTCTCAA	CAAGTTTTAA	ATTAAAAAGA	TTGCAATTTT	TATTATAAAG	33240
CTCCTTAACA	TAATTAAAAT	TAAAATCATT	GCTATACAAA	TCTTTAATTT	CTTCTAAATT	33300
ATTCATCACT	TTTGAAAGTT	CTAAATCAAG	CAATCCAATA	TCATTAAACA	ATTCGCCTTG	33360
TACAGCAACT	AAATTTTTCA	AACTCCAAAA	ATCTGAACAT	AAATAAACTT	TTCGATCCAA	33420
ATCCAAATTT	TCTTTTAATT	TCTTTTGAAG	GGAATGATCC	TTCTCTAAAG	AATTTAAATA	33480
TTTAATCTGA	GAAGATAATT	GATCGTTTAG	AGAAGACATT	TCAATTTCCA	AGCCCAAATA	33540
TCTCTCATTA	GACGCTATTG	CCTGATTACA	TAAAGAAATA	GCTCTTCTAA	TATTTTCAAG	33600
ATCAATTTCT	AATCGATCAA	TATCAACCAA	ATCCAAATAA	CCTTTAAGGG	ATTTACACTC	33660
ACTCTCATCA	TAATCAAGAA	TTGATTTTTT	ATAAGATTCA	GAATTCAACA	ATTTATCTAT	33720
ATTAAACTTT	GTGCGCTCAA	AATCACTTTT	TAAATAAAAT	TCCAAATTAT	CATATTTTTT	33780
CAAATTAAAA	ATATTATCAA	TTATTGCAGC	TTTCTCTTTG	GGAGTTGACG	TTAAAAATTC	33840
TTGAAAATTG	CCTTGTTGTA	AAATTACAGT	TTGACAAAAT	TGGTTAAAGT	CTAATCGACA	33900
AAGACTTTTA	ATATGCTCTA	AAACATCGGT	TCGACCCTCT	ATAATCCTAT	TATCAAAAAA	33960
ACAATTAAGC	AACATGCTCT	TAGGAGTCTC	TATATTTTTT	ACATTAAGCT	CAACAAAAGA	34020
TTCATAAATC	TTCCCAGAAA	TAGTAAACGT	TAATTTAACA	TAAGCGCTAG	TCTCGCCTTT	34080
TGATATAATA	TCTACAATTT	TTTTTCCAAG	TCTGTAAACA	CGAGCATATA	GTGCCAAAGT	34140
TATGCAATCT	AAAATGGTGC	TTTTGCCTGA	TCCAGTATTA	CCAGAAATTA	AAAAAATGCC	34200
CGATTGTCTT	AAAAGAAGTG	TATCGAAATT	CAACTCATGT	TCGCCTTTAT	AAGAAGCAAT	34260
ATTTTTAAAT	ATGAGCTTAT	TTATCCTCAT	ATTCACCCAA	ATATCCGTTA	GCTAAAACCT	34320
CGTTAAAAAG	AGAAATAAGC	TCTTCTTCCT	TAAATTTGAT	ATCCCTAATA	ACACCGTTCT	34380
CAAAATCCCA	TCTCAACTTT	TTCTCAAAAA	AATATTTTTT	ATCCATTTCA	AGCACTTCAA	34440
GTTCTCCAAT	AAAGTTGGAA	TCGTCTTGCA	AATCCTGACT	TGAGGGTAAA	GAATAGGAAA	34500
TAGAACTAA	ATTCATAAAA	TTAAGCCTTG	CTAAATCATA	AATAGACTCC	TCAGCGCTAG	34560

TATCAACTGC	CTCATTAAGT	TCAATTTTCA	AATAAATAGT	AAAAGACTCT	TCTTTTTTGG	34620
AATTAGCCAA	AAAATCAAGA	ACTTCATTTA	AAGAACCTTT	AGCAAAGATT	AATTTATTGA	34680
AAATCGGCAC	CGGAAATGCT	TCTTGCAAGA	TTAATTTATT	GTCATTAAAA	TGTAAAACGT	34740
TTATGTATTT	ATCACAGGTC	TCATTAAATG	AATATTGCAT	AGGAGATCCT	GAATAAACAA	34800
TATTATCTCT	TAGTTTCATG	AACTTATGAA	TATGCCCAAG	AGCAACATAA	GAAAAACCAT	34860
TTCCAAAAAC	ATTAAAAGGG	ATAATATAAC	TACCTCCCAA	GGTGTCAATC	TTTTTACTGC	34920
TGCCAAAAAA	AGAATGCGCC	ATTAATATCT	TAGGAATTCC	TTTATACTTG	TTTTCTAAAA	34980
AATTAGATAA	ATTTGATATT	TTTTCTCTGT	AGGCATTTTC	TAAATTTTCA	AGAAATAGTT	35040
TGCTGGAATA	CTGATCTTCC	AGTCCAAAAA	TATTGTCAAA	ATTTTGACCT	AAAATAAGCC	35100
TTTCATTTAT	ATGCGGAAGA	CAAACAACAA	TAAACTTAAG	ATTTCCATTA	TCTTTTAATA	35160
AAACTATTTG	CTCATCAGAA	TCATATTCAG	TTATTAAAAA	AAAATTAAAC	CGTGAGAGAA	35220
GTTTTTTTATT	TATACTCAAA	TAATCCTTTT	TGTCATGATT	TCCAGAAATA	ACCACACACC	35280
ATTTACAAGA	AGTAAAAGAA	AGTTCATAAA	AAAATTATT	CACTAATCTT	TGCTCTTCAA	35340
ACCCAGGCCT	TTTGGAATCA	TAAACATCCC	CGGCAACAAG	TAAAAGATCT	ATATTTTCTT	35400
TTTTAATAAA	TTCTAAAAGA	AAATATAAAA	AATTTTCTG	CTCCTTAAGA	ATTGAAAAAT	35460
TTTCAATTTT	TTTTCCAATG	TGCCAATCTG	AAGTATGCAG	AATTTTATAA	TTGCTCACAA	35520
AATACTCTCA	CTTTTTTTAA	TTCTTAAATT	ATATTTATAT	ATTATAATAC	AATATATAAA	35580
CATAGGGAAT	TTATGCAAAA	TAAAAAGTTG	ATAATAGTTG	AATCGCCAAC	AAAAGCCAAA	35640
ACAATAAAGA	AGTTTCTCGA	TGAATCATTT	CTAGTAGAAG	CATGCATTGG	ACATGTAGTA	35700
GATCTACCAA	ACAACGCAAA	AGAAATCCCA	AAAGAATATA	AAAAATACGA	ATGGGCAAAT	35760
ATTTCTATAG	ATTATAACAA	TGGATTTAAT	CCAATTTACA	TTATTCCCAG	CAATAAAAAA	35820
CCAATTGTAT	CAAAACTAAA	AAAATTAGTA	AAAACAATAA	ATGAAATATA	TCTTGCAACC	35880
GACCAAGACA	GAGAAGGAGA	AACTATAGCA	TTTCACTTAA	AAGAAGTATT	AAAAATCAAA	35940
AACTACAAAC	GGATGATATT	TCATGAAATC	ACAGAAACCG	CAATAACTGA	ATCACTAAAA	36000
AATACTAGAA	ATATAGACAT	GAACCTTGTT	AATGCCGGGG	AAGCTAGAAG	AATATTGGAC	36060
CGACTATACG	GGTATACAAT	CTCTCCACTA	CTTTGGAAAA	AAGTAGCTTA	TGGACTTTCT	36120
GCTGGGCGAG	TACAATCTGT	TGGATTAAAA	TTATTAATAG	AGAAAGAAAA	AACTAGAATA	36180
AATTTCAAAA	AGGCAAATTA	TTATTCAATT	TTACTTCAAT	GTAAACACGA	GAAAAAAAAC	36240
TTGTTGCTTG	AAGCAAAATT	AGAAGAAATT	GACGGCAAAA	ATATAGCAGA	GGGTAAAGAC	36300
TTTGTAATG	AAACTGGAAA	ACTTAAAAAT	ATTGCCAAAA	CAACAATAAT	AACCCAAGAT	36360

TTAATGATAG	AGCTTGAAAA	AGAATTAAAA	AATGGACAAA	AAATTGAATT	AATTTCAATA	36420
GAAACTAAAA	AAATAAAAAAT	ACCTCCTCCA	AAGCCATTTA	CCACCTCTAC	ACTTCAACAA	36480
GAAATAAATA	AGCGTCTTAA	AATTGGAACA	AAGCAAATCA	TGCAACACGC	TCAAAAACCTT	36540
TACGAACACG	GATACATTAC	CTATATGAGA	ACAGACTCTC	ATAATATTGC	TAAAATTGCA	36600
AAAGATAAAA	TAACAAAAAT	AATAAAAAAT	AAATATGGGA	AAGAGTATAT	AGAGGAAAAA	36660
GATAGAATTT	ATGAAAAAGA	AAAAATGGCT	CAAAATGCAC	ATGAGGCAAT	AAGGCCTTCT	36720
GAAATATTTA	TTCCAAATGA	AACCATAGAA	ATAGAAAGCA	AAACCGCTAA	AGAAATTTAC	36780
AAAATAATAT	GGGATAGAAC	CATTATTTCT	GGAATGAAAG	ATGCAATAAA	AGAAAATATA	36840
AAACTGACTT	TTAAATATAA	AACTTAATT	TTCAGATCAA	GTTTTACAAA	AATAATTTTT	36900
GATGGATTTT	TTAAACACAC	TAAAGAACAA	GATGAACATC	TTAACATAAA	TTTGTACTTA	36960
ATTAAAAAGG	GAGATACATT	TTCCATAGTT	AAAATGAAAA	CAAGTGAGCA	CGAAACAAAG	37020
GCTCCATTTA	GATACACAGA	AGCGTCTCTT	GTGCAAAAAA	TGGAAAAAGA	AGGAATAGGT	37080
CGTCCCTCGA	CCTATTCTAC	AATTATATCA	ACACTTTTAG	AAAGAGAATA	TGCATTCAAA	37140
CTTAACAACA	CATTAATGCC	AACTATAAAA	GGCGCTGCTG	TAATAAATCT	TCTTGAGAAA	37200
TATTTTCCAG	TACTCATTGA	ACTAAATTTC	ACCTCTAATA	TGGAAGAAAA	ATTAGACAAA	37260
ATAGCAATAG	GAAAACTAGA	TAAAATAAAA	TATCTAAGTA	AATTTTATAA	TGGCAAAAAA	37320
GGACTAAAAG	ATACAGTAAT	GCAACTAGAG	CCTAAAATTG	ATTCCTCTGA	ATTTAGAACC	37380
GTTATTGAAA	GTCAAAAAAT	AGAAAATAAA	AATAGCATT	ATTACACAAT	AAACATTGGT	37440
AAATATGGGC	CTTATTTGAT	ATTCAAAGGA	CATAATTACT	CAATTAATGC	AAAAACTCCA	37500
TTAGAAAATT	TGTACAAAAA	AGATGAAATA	GAAAAAATAA	TAAATGAAAA	AGAGCTAAAA	37560
CCCAATATAC	TTGGGGTTGA	TCCTTTAACA	GGACTTAATG	TGATCTTTAA	AAATACAATT	37620
TACGGAAACA	TTGTTCAACT	TGGAGAAGAT	ACCCATGCCC	CTCAAGAATA	TACAAAAAAA	37680
GGAAAACCTA	AAAAATTAAA	AATAATAAAA	GCAAAAAAAG	CATCAACTAA	AAAAATTGAC	37740
CCTGAAAACA	TAACATTGGA	GCTTGCTTTA	AAATTGCTCT	CACTGCCAAA	ACCAATTGGC	37800
AAACATCCCC	AAACCAATGA	ACAAATCATT	GCTGCAACTG	GTGTTTTTGG	GGATTATATT	37860
AAAAGTAAA	GCGGAAGCAT	TGCTTGCTCG	CTAAAAAAG	ATTTAAAAGC	ATATGACATA	37920
ACACTAGACA	AGGCCATCAG	CCTACTCAAC	GAAAGAGCCA	ATAAAGTGGG	TATAATCGTT	37980
AAAACAATCA	CATTTTCTAA	AAACAAAATT	GGCAACAAAA	TATATATTTA	CAAAAAAATC	38040
GACAAATTTT	ATGCTAAAAT	TAAAAGAAAG	AAGATTGATT	TACCTGATAA	CATTAATCTT	38100

GAAGAAATAA	ATGAGAAATA	TGTATTCAGC	TTGTTATAAA	TATGAATGAT	TTCAAACCTCC	38160
CAATTTATAA	ATACAAAGAT	GAATTAATTA	AAGTACTAAA	AAACCACAAT	GTTTTAATTG	38220
TAGAAAGTCC	AACAGGTAGC	GGAAAAACCA	CCCAACTACC	AAGAATAATA	TATGAAGCGG	38280
GTTTTGCAAA	ATTAGGAAAA	ATTGGAGTAA	CTCAACCAAG	AAGAATAGCT	ACAGTATCAA	38340
TAGCTGAATA	TATTGCCAAG	CATATTGGCG	TAAATGTTGG	AGAAGAAGTT	GGCTATAAGA	38400
TAAGATTTGA	AGAAATTACA	AGCCCCAAAA	CCAAAATCAA	ATTAATGACT	GACGGAGTGC	38460
TTCTGCAAGA	GCTAAAAAAA	GATACACTGC	TTTATGAATA	TGATGTAATA	ATAATAGACG	38520
AAGCACACGA	AAGAAGTTTA	AACATTGATT	TTATATTGGG	TCTTATCAAA	GACATTTCAA	38580
GGAAAAGGGA	TGATTTTAAA	ATCATAGTTT	CGTCTGCTAC	AATAAACACA	AAAATATTTT	38640
CAAAATATTT	TAATAATGCA	CCGGTTGTTA	GTATTGAAAC	TATCACTTAC	CCAGTACAAA	38700
TAATATACAA	TCCTCCTCTT	TTAAACACAT	CAAAGGAAT	GATATTAAAA	ATAAAAGAAA	38760
TTGTCTTAAA	CGTAATAAAA	GAAAAAAAAG	CGGGAGATAT	TCTTATATTT	TTATCTGGAG	38820
AGAAAGAAAT	AAAAGAACT	ATAAAGAAT	TACAAGAATT	AAACTCAAAA	AAAAATTTAA	38880
TAATATTTCC	TTTATACGGC	AGAATGCCCA	AAGAAGCTCA	AGAGCAAATA	TTTATGACTA	38940
CTCCTAAAAA	TAAAAGAAAA	ATAATAGTGT	CAACAAACAT	AGCAGAACT	TCAATCACAA	39000
TTGAAAATAT	TAAAATAGTA	ATAGATAGTG	GAAAAGTTAA	AACAAATAAA	TTCCAAACAA	39060
AAACTCATAC	CTATTCGCTC	CAGGAAGTTC	CAATTTCAAA	ATCATCAGCA	ACTCAAAGAG	39120
CTGGTCGAGC	AGGAAGACTT	TCAAAAGGAA	CTTGCTACAG	ACTTTACAAA	AGAGAAGATT	39180
ATCAATTAAG	AGAAGATTAT	CAAAAAGAAG	AAATATATAG	AACAGACCTA	TCTGAGGTAG	39240
TGTTGAGAAT	GGCAGATATT	GGAATTAGAG	ATTTTACCCA	CTTTGACTTT	ATCTCAAAAC	39300
CATCAACGCA	TTCGATTCAA	ACTGCAAGCA	AAATATTAAA	ATCTCTGGAT	GCTATAAACA	39360
ATAAAAACGA	ACTTACAGAA	ATTGGGAAAT	ATATGATACT	ATTCCCATTA	ATACCAGCAC	39420
ATTCAAGAGC	ATTAGTCGAA	GCAATGATAA	ATTACCCACA	AGCGATCTAT	CAAACCACAA	39480
TAGGTCTATC	ATTTTATATC	ACAAGTGGAA	TTTTTCTACT	ACCCCAAAAT	GAAGAAATGG	39540
AAGCTAGACA	AGCTCACTTA	AAATATAAAA	ATCCAATGGG	AGATTTAATT	GGGTTTGTTA	39600
ATATCTTTGA	AGATTTTAAA	AAAGCTCTAA	ATAAAGAAGC	TTTCACAAAG	GAAAAATTAT	39660
TAGATCTACA	AGGACTTGAA	GAGATAGCAA	ATGTGCAAAT	GCAGCTTGAA	AACATTATTA	39720
GCAAATTAAA	TATACCAATA	ATACAAAAAG	GTGTTTTTGA	CAACGAAGGA	TATTTAAAAA	39780
CAATAATGAG	AGGAATGAGG	GATTATATTT	GCTTTAAAC	TTCAAAAAAG	AAATATAAAA	39840
CCATCAAGGC	TCAAAACGTA	ATAATTCATC	CTGGATCACT	TATTAGCACC	GATTCTGTGA	39900

AATATTTTGT	TGCAGGAGAA	ATTATAGAAA	CTACAAAAT	GTATGCAAGA	TCTATTGGTG	39960
TCTTAAAAAA	AGAATGGATT	GATGACATTA	TCCTTAATGA	AGAGTTTAAA	CATAACGACA	40020
TATCTAGCAA	AGAGAACCAA	ATAACAAATA	CCGGGCAGAC	AAAAATTATC	AATGAAATCA	40080
AAATAGGGAA	AAAAATTTTC	AAAGCGGAAT	ACAAAAATAA	CATTTATGTA	ATAAAAATTA	40140
ACCTAGAAAC	GCTAAAAGAA	ATAATTTTTA	AAAACGAACT	AAACAATCAA	AATAATGAAG	40200
ATCTCAAAAA	AATTAAAAATA	CAATTGATGC	ATAAAAATAT	AACGGTTTTT	AACAACAAAA	40260
AATTTTTAGA	AACTATAGAA	ATAGTCAAAA	ACATGGGAAA	AGATTGGCAT	TGTATAAAAA	40320
AATATGAAAC	AAAGAATGTA	AACATTGACG	AACCTGAAAA	AATGAAAAAT	CTTTTAGAAT	40380
GCACAATGCA	ATTTATAAGC	TTTCCCCCA	AAAAAACGC	TCTTTTTTTA	TCGTTGGAAA	40440
CAGATTATTC	TGGAAATTTT	AGACTAAAAC	CCAAACAAAA	TTTCATAATG	GCAATAGAAG	40500
AATCTATAGA	AAGCATAAAA	AGCCTTATAG	AAAACAAAGA	ATACATACAA	AAGTTACATT	40560
TTATAAAAAA	ATTAATAAAT	AAGGTTTACA	AAAAATTAAA	TTACTTTTTT	TAAAACTAA	40620
ACTTTGAAAG	CCTTGTTATA	ATATAAAATA	TAATAATCAA	ATATTATTCA	AAGTTAACAG	40680
CAATGAAGTT	TATAATAAAT	TATGAACCTG	CTATCCTTTT	TTTATGTTTT	ATTATTTTTA	40740
TTAATTTTTT	CTTTTGAATT	ACAGAGTAAT	AATAAAGAAA	ATATAGAAAA	TTTAATAAAG	40800
CTACATATGC	TTTATGATTT	AACCAATAAC	CTGTCAAAAG	AATTAGAAAC	AATAAATAAA	40860
ATTAAAAATT	TTGACTTAGA	ACAACATTAT	CTGCTAATTA	CAAAATATTA	TCTAAAAATA	40920
AAAAAATATA	AAGAAGCTAA	TGATTTTTTT	AAAAAATAA	ACCAAAAAAA	GATCAAAAAT	40980
CAAAAAATAA	AAAACGAAAT	CATTTGCTA	AAATTAAGAA	TAAATGAAGA	TAATATTAAT	41040
GAAGAAGAAA	TCAAAAAAAT	TTTAAATAAC	GAAAAAATA	TAGATGTCAA	AATAATTTAT	41100
CAAATATTCA	GTCTTATAAA	ATTTAAAAAT	AAAAAATTAG	CAAATAAAAT	TAAAAACATA	41160
ATACTAACAA	ACTATCCCAA	AAGCATTTAT	TCTTATAAAA	TAAAAAGAAA	TGAATAAAAA	41220
AATATTAACA	CTGCTAGTAT	TGATTTTAAG	TATTTTCATCA	GTACTAATGC	TGTCCAAATC	41280
AATCACCAAA	AAATCCAAAT	ACAAAATTAT	TAGGGATTAT	TTCATAAACA	GCAATTATGT	41340
TCTGGTGAAA	ATTGAAAATA	AAGATCTAAA	ATTTACCATA	TCAAAACCTA	TTTACGACAA	41400
AAAGCTAAAT	AATTACTTCT	TTAAAGGCCA	AACAACAAGC	CATTTCTTAA	TTTCTAACAA	41460
TGTTGACATT	GCAATTAACA	CAAGTCCATA	CGAAGTTAAA	CAAAACATGT	TTTTCCCCAA	41520
AGGACTATAC	ATATATAATA	AAAAAATGAT	TTCAAAACAA	ATAAATAACT	ACGGAGAGAT	41580
TGTAATAAAG	CACAACAAAA	TTATATTAAA	TCCCAAGGAA	GACGAAATAG	AAAACGCGA	41640

TTATGGATTT	AGCGGATTTT	TTGTTTTAAT	CAAAAACGGA	AAGTATAAAA	AAAATTTTAA	41700
AGAAACAAGG	CACCCAAGAA	CAATAATAGG	AACTGATAAA	AATAACAAGC	ATTTATTTCT	41760
TGTTACAATA	GAAGGAAGGG	GTGTCAATAA	TAGCAAAGGG	GCCTCTCTTA	ATGAAGCTAT	41820
TGATTTTGCA	TTAAGCTACG	GCATGACTAA	CGCTATTAAT	CTAGACGGGG	GGGGCTCAAG	41880
CACTCTTGTT	GTAAAATCAA	ATAACGCTCC	TTACAAATTA	AACTTCACAG	CAAACATCTT	41940
TGGACAGGAA	AGACCTGTCC	CATTTTCATTT	AGGAATAAAA	CTTCCTAATT	GAAAAATCTC	42000
CAACCGATAT	TAAATCCAAG	CATAATCTCA	GTTGTTAACC	CAGAAAAAAT	TTTATAATTA	42060
GAAATGGAG	AAATAGAAAG	CATAAACAAA	GGCCTAATAT	ATAAACCATC	AAGATCGGGA	42120
ATAAAAAGAT	CAGCAGCAAG	CCCCATGCTT	AAAAAAAAGA	GATTGAAGTT	ATTAGAGCTT	42180
AAATCAAAAG	CATATTTAAA	CCCAATTAAT	GGGAAAAGCA	TTCTAACCTG	CTCTTTGAAA	42240
ACCATTGGAT	ATGTTCCATA	AAGCCCAAGC	GAGAAATATC	TCCCATTGTG	AGTAACAACA	42300
AAAGCCTCTT	TGTAAGACAT	TTCAAAAAGT	ACATAATTTG	CATCAAAAAA	TAAATTCAAA	42360
TTAATCCCAT	GATCTGCTCT	GGTAAAATTT	GGAGCAAATT	TAGTGGCGCC	TGTTTTATCA	42420
GTATAATTAG	TAAATTGATA	AGAAAAACCT	CCACCAAAAG	AAAGTGGATA	AGAAACAATT	42480
AAATTATTAG	AAGAGATGAG	AAATAAAATA	AAAAAAGAT	ATTTCTTCAT	TAACAATCCT	42540
TAAAAATTCT	AAAAAATACT	ATATTATTAT	AGTAACACAC	TAAAGTAGTA	TATAAAAAAT	42600
CTGGGAAATT	ATGAATACAA	AAACATTATA	TTTAATATCC	TTAATTCTTT	TAGCTTGCAA	42660
TAAAAATAAC	AAAATTCCCTC	TCATTCAAAA	ATTAGATTTG	CCCAAAAGCA	GCATTCTTGG	42720
CTTTAGCAAT	AAAATGGGCA	TAATAATAAA	AGATTATGCT	TTTCTTAGTA	AAAGCACTAA	42780
GAAAAATAGC	GAATTGGATT	ATGATTACGC	AATTCTACTC	AGAAAAGACG	AAGTCGTAAA	42840
AATTGAAAAA	ACACTAGAAA	AAACAGAGCG	CTATGGAATT	GAAGGAAATT	GGATCCTAGT	42900
CAATTACAAG	GGAACATAAA	GATACATCTT	TAGCAAAGAC	ATCAATATAG	TCACAATTTT	42960
AATAATTGAT	CATTCTAAAT	AGCTTTACTA	CATAACCGGA	CAAAAGTCCG	ATCAATGTAA	43020
TAAATTACTT	ATTTTTTTTC	TTATGTCTAT	TTTTTCTTCT	TTTTTTCTTC	CTTTTATGGG	43080
TAGAAATTTT	TTTTAATTTT	CTTTTTCTTC	CGCAAGGCAC	TCCCTAAATC	TCCTTATCAA	43140
CCTTAAAAAT	TAAATTCAAA	ATGTTATCTA	AATCATCCTC	TGGATGAAGC	CATAAAACAT	43200
CAGAAATTTT	GGCAAAAAG	GTCATTTGCC	TTTTTGCATA	TAAAAACGAA	TTTTTGTTTA	43260
TTAAACCTAT	TATATCATTT	AAACCATAGC	AAGGTCTACT	TTTCATAAC	AAAAACTCAT	43320
TATAGCCTAT	TCCTTTAAAA	GCCGGAGTAT	TTTCATTGTA	ACCCTTGCTA	AATAAACCTT	43380
TAATCTCGGA	AAGTAGTCCA	CTATTAAGCA	TTTCATTAAT	TCTTATTGAT	ATTCTGGTTT	43440

TCAAATCTTC	AAAAGATCTC	TTAAGGCCTA	TAATCACAAT	ATTTTTAAAT	TCGCTACTTT	43500
GTTTTTTTTG	AAATTGGCTA	ATAGGAATTC	CTGTTTGATA	GTAAACCTCA	AGCGATCTTT	43560
TAATGCGATA	AATATCATTC	TTATTTAACA	TATTAAATCT	GATGGGATCT	ACATTTTTTA	43620
ATTCTTTTAA	AAGATAAGAT	TTACCCTTAA	GCTCTAAAAG	ATTGTTTACA	TAAATTCTTA	43680
TTTTAGAAGT	AACCAAGGGT	GTTGAAGGAA	ATCCATCCTT	TAAATGCTTA	AAATAAAAAG	43740
CAGTACCTCC	TACAAATATA	GGAATTTTTT	TTTCTGTCT	TATTTCTTTT	ACTATTTTTA	43800
AAGCTTGTTT	GTAAAAAATT	CCAATAGTAT	AATCCTTTTC	GGGATCTAAA	AAATCTACTA	43860
AATGATGCTT	TATATGTTTC	ATTAAATTTT	TACTTGGCTT	TGAAGAAGCT	ATATTAAACT	43920
CTTTATAAAC	TTGAATAGAG	TCAACATTAA	TAATTTCTGC	TTTATTTTTT	GGAAAATGAA	43980
ATAAAATATT	GCTTTTGCCC	ACAGCTGTAG	GGCCAAAAAT	AAAACTACT	CTATCTTCCT	44040
TCAATTGAAT	ATCTAAATTT	ACCAGTCAAA	ACGTCATTAA	AAGCTTGCCC	TAATATTTTA	44100
TCATCAAAAA	TAGCGTGTTT	TGCTAAAGAA	ATATTGTCTA	TAATTTGCTC	AGTGCGCATT	44160
ATAGTTGCAA	CAACAAGTTC	ATAAAAAATT	CCATCAAAAT	CTTGATTTTT	TTTTAAAGGC	44220
ACTTTAGTCA	TTAATCTCC	TTATAAACCA	AACATATTAA	ATAAGTTTAG	CTATTTTTTT	44280
TCAAGTTTCT	TTTATCAATT	TCATCTCTAA	TTAATTCAAA	AACTTCATTT	GGATTAATAT	44340
TTGTAACATC	AATCACTAAA	TCGGTCTCTG	AAAAATAATC	ATCAATATCT	ATATTGTAAA	44400
TAGCTAAATA	TCTTTTTTTG	TCATTTTCAT	CTCTAATAAA	AGTACTGCTT	AAAACGTCAG	44460
AATACATGCC	CCCCTCTCTA	GTCATTATTC	TCTCAGCTCT	AACTTCCATT	TTAGCATAAA	44520
GATATATTTT	TAAATCAGCA	CTCTTAGAAA	TCCAAATAGC	AAGACGAGAT	GCAAGCACTG	44580
TATTATTTTT	TCTAGAAAGC	ACAGACAATC	TATTATCAAG	GTATTTATCC	CAATAATAAT	44640
CATTTCTGCC	TATTATCTCT	TTTTCATAAA	ACTCTGAAAA	AGGAATATTA	TGCTCTCTTG	44700
CAATATCATG	AAAAGTATAA	TTAATAAACT	CAAGACCGTA	ATGTTTGGCA	ATCATCCCGC	44760
TTACAGTAGT	ATTGCCACAA	CCACTCTTAC	CAGAAAGTGC	TATTTTCATT	CAACATTCCT	44820
TATTTGCTCT	TTTATTTTTT	CTAAATTTAA	CTTCATATTC	AAAATCAAAT	TCTTAATATC	44880
AAGATCAACC	GCCTTATTAC	TCATGGTTGT	TATTTCTCTG	TGCATTTCTT	GAGAAATAAA	44940
TTCAAGAGCT	TTACCACATA	TCTCATATTC	AAGATTTTTA	TAAAAAGTTT	CTATATGAGA	45000
ATCTAAGCGC	ATGATCTCTT	CATTAATGTC	TAGACGAATT	GCCATTTTAG	CTGCCTCTTC	45060
TGCAATATTT	AAATCTCTAA	ATTCATCCAT	TAATTTAGAA	ATATTTTCTT	TAATACTTGC	45120
AAATAATTTG	ACATTTATAT	CACTGCAAGC	ATCTTTAACA	ATTTTAAGGT	CCCGCTCTAT	45180

TAACACAAGG	GTTGACACTA	TGTCTGACTT	GGTATTTTCT	CCTTCAAAAC	TTCTTCCATT	45240
ATTGTAATGT	AATAAAGCTT	CTTCTAGAAC	ACCTTTTAAAC	AACCCATAAA	TCTCTTCTTG	45300
ATGTTCACTA	TCCTCATCAA	TTATCAAAGC	TCCTTTTAAAT	GATAAAAAAT	CGCCCCAACT	45360
TAGTTCGTTT	TTAATATTAA	GATTGGTATG	TGCCAAAGAA	TCTCTAAGCC	TAGAAATAGC	45420
CTCAATATAA	TTGGGATTAA	TCGTAAAATT	CACACTAGGA	ACCAATTCTT	TATATCCTAC	45480
ATTTAAAAAA	ACATTGCCTC	TGCTAATATA	TTTTGAAATC	AAATTTCTTA	TATCAAGATC	45540
ATAGCCAGAA	AAAATTTCTG	GTAACCTAAA	TTTAAATTCT	AAAACTTTC	CATTATAAGA	45600
TTTCAAATTA	ACACTAAACA	TATAGTTACC	AATTATCTTT	TCCAAATAAA	AAAATCCCGT	45660
CATGCTTTTC	AATATAACAC	CCTTACAGAA	AATCGTGTA	AATAAAATTA	TTTCCAGCGC	45720
CCATTGTAAT	AAACAAGTCT	CCAGATATTA	ATAAACTTTT	TATAAAATTA	ATAGAGTCTT	45780
TAACATCCTT	AAAAAAATAA	GTATTCTTAT	TTATTTTTTT	AATATTTAAA	AACAATTTAA	45840
CAGAAAGTTC	ATCTGGATTA	AAATTTTCCC	TATTTGAAAG	ATATATATTG	TGCAAAATTA	45900
ATATATCGGC	AGCACTTAGA	ACTTCAACAA	AATCGGCAAA	AAATTCCTTT	GTTCTTGTA	45960
AGGTATGAGG	CATAAAATCC	AAAATTATAC	GTTTATCTTT	ATAAAAATTT	TTAATACCAA	46020
AAAGAGTATT	TTTAATTTCC	CTAGGATGAT	GAGCATAATC	GTCCATGTAA	ATCACTCCAT	46080
TTTCCTCTTT	AACAACTTCA	ACCCTTCTTT	TTATACCGCT	ATAATTTTTT	GCAATTCTCT	46140
TTATTGCTTC	TTCAAAATCA	AAAATTGATT	TCCCATTACT	TTCTAAGAAA	AGATTTAAAG	46200
CCAAAAGCGC	TGCTGAAAAA	TTTAATACAT	TATGAAATAA	AACAGTCTTA	AGCTCAACAT	46260
TTAACAAGCC	TAAAAAAGAA	AAACAAAAAT	ATTCACTCCT	AAC TGCAATA	TTACTTATTT	46320
GAAATCAGA	TAAATCTCCA	GACCCATAGC	TAAAAATACT	TATATCTTTT	CTGTTGATTT	46380
GCCTTTTAAAT	TTTAAGCAAA	TTATTATCAT	CGGAATTAAT	TATCAATATT	CCATTTTCTT	46440
TTAAATTATT	AATATACTGT	AAAAAAGCCT	CTTCAAGAGC	CTCATAATTT	TTAAAAAAAT	46500
CAACATGCTC	GTAGTCAACA	TTGGTTAAAA	TAAGCATATT	AGGGCTAAAA	TTCAAAAAAT	46560
GTTTCTTATA	TTCAACAAGT	TCAACAATAA	AAATATTGCT	AATACCTGCT	ATTGCAGAAT	46620
TATCTTTAAA	ATCTTTAACA	CTTGACCCCA	CAATAACATT	GGGATTTAAT	CCTAATTTAT	46680
TAAAAAGAAC	ACCTAAAAAC	GCCGTAGTGG	TAGTTTACC	ATGAGAACCT	GCAATTCCAA	46740
TGCTATAGTA	CTTTCTAGAA	AGCTCTCCAA	GAGCCTCAGG	ATAAGATAAA	ATAGGTATAT	46800
TTAATTCTTT	TGCCTCAAGT	AAAACCTGCA	AACCATCCTT	ATTATAGGCT	GAAGAATATA	46860
CTATTAAATC	AAAAGACCTA	TCAAGCTGTT	TTAATGAAAA	CTCATAAATA	TTATCATAAT	46920
AAGATATTTT	ATTATTACTT	AAAATTTTCA	CGGTATAAAA	TTTATCAGAA	ACATCTACCC	46980

CTTCTACACA	ATACCCCTTTT	GAATTTAAAA	AACAAGCCAG	AGAACAAGCC	CCACTTCCCT	47040
TTATTCTTAC	AAAAAAAATA	TTATTCAAAT	CGTCAAAATC	AACCTTCATA	GCTCTCTTCT	47100
AAATAATCTT	TTTTGTCATA	AAATTTATGT	ATAACTATAT	CGATAGATAA	TATATTAATT	47160
ATTGTATACT	TAAGTACGTT	CGTAATAATA	TAGGTTCTAA	GCATTTCTGT	ATTATTAAGC	47220
AAGTTATTTT	CAAGAGGAAT	ATTTAAAAAC	AGTTTAAAAA	GGTAATTGTT	TCCATCTTTT	47280
TTGATTTTAA	GATCATAAAC	AATATAATTT	CTATCATACT	CAGAAACACA	ATGTTCAATA	47340
ATTTGCCTTA	CAGCTCTCTT	AGAAATAGAT	AAAACCCCTT	CTTCATAAAA	ATGGGGCCTT	47400
ACAACAGATC	GAATATAATT	TTTCTTCCTA	GCAAAGAACC	ATCCGCTTTT	AAAAAAAAC	47460
TTAATTGAGT	TTAATAGCAA	ATTGGGCCTA	ATAGACGTTA	TTTCAAAGC	AGCTGCTGGT	47520
ACAACATGCT	CCCCCATTTG	CCTTGAAATT	CTTGCTTTTT	CTATTTCTTG	CCTGGTTGAA	47580
ACATCTGTTA	TGTAAATAAT	CTTAAAAACA	TTTGGCAACA	AAAGTTTGA	AATTATTTTA	47640
TCTATCATTT	TAAGACTTGT	TCCTAATATT	AATATTTTAT	TAAATTCTTC	TTTTGCAAGC	47700
ACTTCAAGCA	TTTCTCTGCA	ATGAGCATCA	TCTTCAAATA	CAGATCGCCT	TACCGCTTTA	47760
AAAACATTAT	CTTCAAACCT	AGCAGAACTT	CCGGCAATAA	TTTTCATATT	TTTAATTAAA	47820
ACACCATCAT	CAATTATTAA	AGGTATTGAA	TATTTATCTG	CTACCAAATG	CGATCTAAAG	47880
CTCTTGCCAG	TTCCAGCAGA	TCCTACTAAT	GCATACACCT	TAACCCTAAC	AAATTTATGC	47940
TTAATGCTAA	TTGCAAAGTC	TTTAACCTTG	CTTAATATTT	TTTTTAAAAA	AAAATCACTT	48000
GAAAAATTCT	GAAAATTCAT	AGACATCAAA	AATATCTTTT	ACTTGCTTTA	AAAATCTAAA	48060
CTAGGCTTAG	AGCAAAATTC	ATTCTTTTTA	AAGAATGTTT	CATTAAATTT	TACCAAAAAA	48120
GCATGTAAAA	AGTAATCGCT	TTTCTTGAAT	TTATTACAAT	ATTTCTTGTC	ATTAATTAAA	48180
GGATGATTGT	TAAAGGAACA	CTGAGACCTT	ATTTGATGGG	TAAAGCCTGT	TTCAATAACA	48240
ATCTCGACAA	GAGTAGCTCT	TTTGCAAGAT	AATATTGGAT	TAACCTTTGT	AATTGCATTA	48300
ACAAAATTTT	TATCTTCTAA	AACAAAAGTT	TTTCTCAACC	TTTTATTTCT	AAATAAATGA	48360
TTTTTATAAA	CAACAGGAGA	CTTAACCTCG	CCTAAAAGTA	TTGCAAAATA	TTTTTTAATT	48420
ATAGATCCAC	CACTAAATGC	CTCACTTAGC	TTTCTTGACG	TATTTATATT	TTTTGCAAAA	48480
ATAATAATAC	CAGAAGTATT	TCTGTCAAGC	CTGTGAACTG	CCGAAGGCTT	AAAGCTTAGG	48540
GATCTTAAAT	TTTGACTTAA	AAGATAAGAA	TTCATAAAA	AATCAAGAGA	ATTTTTACCT	48600
CCATGAACATA	AAATACCTTT	TTGCTTATCT	AAAACAAGTA	AGTCACTGTC	TTCATAAATT	48660
ATTCCTTTTC	GAATATATTG	AAAATCAATA	TTGCTTTTAA	AGCATTTATC	CGTGGTTAAG	48720

TTCAAATTTT	GGGCTAAAGA	TTTGTACAAA	TAAATTTTAT	CACCTTTGCA	AACTCTGCAT	48780
GAAAAATGTG	ATTTTAAACC	ATTTAGCCTA	ATGTCACCTT	TTCTAATATG	TTTTATTATA	48840
CTCGCTTTAG	AAAAATTTAA	AATTTTAATT	AAAATTGAAT	CTAGTCGCTT	GCCATTATCA	48900
TTAGCAAGCA	CTTCTAAAAA	AATATATTTA	TCCAAACGCA	AAAAATACCC	CTAACAAACC	48960
TTACTATTTT	TTTTACAAAA	AAAATTAACT	ACTAAAAATG	TAAATATAGA	AACAAAAAAT	49020
GATGGAAAAA	CGGGGTGAAA	AAACCAAATA	TTTAAACCAA	AGAATAAAAT	GGACAAATAA	49080
AATATTAACC	CTAAAAACAT	AGAAGCAAAA	GCCGCTATTT	TGCTTACAAA	ATTTAAATAA	49140
AGTCCAAAAA	CAATAATAGG	GAAAAACGAA	ACTTCCAAAG	CTCCAAAGGC	AAAAATATTA	49200
ATAAGAATA	AAAAATGCGG	AGGAAAGAGA	GAAAATATAA	GTATTATTAA	AATAAAAAAA	49260
ATATTAGAAA	TCATTATTAT	TCTGCCAATC	TTTACATCTT	CTTTTAAATC	TTCTTTATAA	49320
ATAAATATTG	ACTTTATTAA	AACAGATGTT	ATTAATAGCA	AATTTGAATC	CACTGTAGAC	49380
ATTATTGCAG	ATAAAAGACC	TATAAAAAAC	ATAAAACAAG	AAAAAGGATT	TAAAACTTTT	49440
AAAGCCACAT	TTAAACAAC	TTTATCATTT	GGACTTAAAT	CTGGAAAAAG	AATAATAGCA	49500
AAAAACCCTA	TTAAATGCAT	CAAACAATT	AAAAAGCTAA	TAATAAAAGT	AGAAATGGGA	49560
AGAGAAAATT	TTATAGCATT	CTCATCTTTA	AATGCTATAA	AATTATTAAT	AATCTGAGGC	49620
TGCCCTAGTA	TTCTATTCC	TATTAATATC	CAAAAAGAAA	TTATATATTG	TGGCTTTAAG	49680
TCAGCATTTG	AAGGAAGTAA	AAGGCTTTTA	TCTAAGCTAG	ACGTTGCTGT	TTTGAATAAA	49740
TTATTAATAC	CCCCTCCCAA	ATCTAGCATC	TTGGAAAACA	AAATAACGGA	TGAAACTAGC	49800
ATTAAAAATC	CTTGAATCAA	ATCCGTATAA	GCTACTGCCT	TAAAGCCGCC	AAAAAATACA	49860
TAAATAAAAA	CCAAGAAGGC	AAAAAAAAGTA	AGACCAACTA	CGTAATCAAT	ACCCCAAAAA	49920
ACTTCTATAA	GTTTGGCACC	ACCTATTAAT	TGGGCAGAAA	TCAAAAACAT	TGAAAAAATA	49980
ATCAATACAA	ATCCACTCAT	TAACGCCAAA	AAATCACTTT	CATATCTATG	CCTAATATAA	50040
TCAATAATAT	TAATTGCATT	AATTTTTTTT	GATTTCGCGAT	TTAATCTCTG	ACCAACAATA	50100
ATAAAAAACA	TTAAAGTTGT	AGGAATTTGT	ATGGTAGCTA	ATAATATAAA	AGATAATCCA	50160
TACTTATAAA	CAGCAGAGGG	ACCGGAAATA	AACTACTAG	CACTAATATA	GCTAGAAGAA	50220
AATAACAAAG	CCATAACAAT	AAAATTAATA	TTTCGATTTG	CAAGAAAATA	TTTATTTAAT	50280
AACAAAAACC	TACCTCTATT	TCTTTTTTTT	AGAAAATCTA	AAAATAAAAA	AATATCAATA	50340
ATGTATCAAG	TTTTACTAAT	TACTAAAAATA	AAAAAACAAA	CCAAAAAATA	ATTATACTGG	50400
GGAATAAAAT	TCCTGACAAA	AAAAACCACA	AAGGAATATT	AAATATAGTA	GTTGATGTGT	50460
CAATAAAATA	GGCAAAACAA	AACCACAATA	CAACATAAAA	AACATACAAT	AATATAGCGT	50520

ACAAAATCCC	ATTTCTCATT	TAAACAAC	TTAACAAAAG	GACATCAAAA	TTTATAAATT	50580
CAAATATAAT	GTATATTATA	TAATATATAT	TATATGGATA	AAAATAAACA	TATATTAATT	50640
GGTATATGTG	GGGGCATAGC	CTCTTACAAG	TCAGTTTACA	TAGTTTCCAG	TTTAGTTAAA	50700
TTAGGATACA	AAGTTAAAGT	TATAATGACA	CAAAATGCAA	CTAAATTTAT	TACTCCATTA	50760
ACTTTAGAAA	CCATTTCTAA	GAACAAAATA	ATTACTAATT	TATGGGATTT	AGACCACAAT	50820
GAGGTGGAGC	ATATAAAAAT	TGCAAAATGG	GCACACCTAA	TTCTTGTTAT	TCCTGCTACC	50880
TACAACACAA	TATCTAAAAT	TGCATCAGGA	ATTGCTGATG	ATGCATTAAC	TACAATAATA	50940
TCTGCAAGCA	CGGCTCCTAC	TTATTTTGCA	ATAGCAATGA	ATAATATAAT	GTATTCAAAC	51000
CCTATTTTAA	AAGAAAATAT	AAAAAAGCTT	AAAACCTATA	ATTATAAATT	CATTGAACCT	51060
GATAAAGGAT	TTTGTAGCTTG	CTCATCAAAT	GCTTTAGGGC	GCCTTAAAAA	TGAAGACAAA	51120
ATTATAAAAA	TAATATTGAA	TGAATTTAAT	CAAAAAGACT	ACCTAAAAAA	TAAAAAATA	51180
CTTATAACAG	CATCCAGAAC	TGAAGAATTA	ATAGATCCAA	TTCGCTATTT	CTCAAATACA	51240
TCAACGGGAA	AAATGGGGTT	TTGCTTAGCA	CAAGAGGCTG	TCAAAC TAGG	AGCTCAAGTT	51300
ACAATTATTA	CAGGACCAAC	CAATGAAAAT	GATCCTGAAG	GGGTCAACAT	TATAAAAATA	51360
AAAAC TGCAA	TGGAAATGTA	CAAGGAAGCT	CTCAAAATAT	ATAATAAATT	TGAAATAATA	51420
ATTGGAGCCG	CAGCTGTTGC	CGATTTTAAA	CCCAAACACA	TTTTCAATAG	TAAAATTAAA	51480
AAAAATAAAA	TCAATAGATT	ATATATAAAA	TTAGTAAAAA	ATCCCGACAT	AATCCAACAC	51540
ATAGGACACA	TTAAGCTTAA	AAACCAAATT	GTTATTGGAT	TTTGGCGCTGA	GAATTCTAAA	51600
AATTTAATTC	AAAAAGCTAA	AGAAAAATTA	AAAAAGAAAA	ACTTGGACTT	TATCATTGCA	51660
AATGAAC TTA	AATATTTTGG	TTCAAAATTA	AACAAAGTTT	ATATAATAAA	TAAACAAAGC	51720
ATAAAAGAAC	TGCCAGAAAT	GGAAAAATCA	GAAGTAGCTA	AAGAAATTTT	AAAAATTTTA	51780
TACTAATATG	CTTAATAGTT	TATTAATAAT	CAATAATCTT	TTAAAAGCAT	TTTAATATAT	51840
TCGGAGTTCG	TTTCACTTTT	AATTTTTTTT	AGCTTATCAG	AAAGCGATGA	AGATTTATTA	51900
TAAACAGCTC	CTTTTAATGC	TAAATGCCTT	AAATCAATAT	TTTGCTGTG	AATGATTTTA	51960
GAATAAAAAAT	TATCAAAATT	ACCCTTTTTT	CCAGCCAACA	TTGAAGCAAC	GCCCCTTAAA	52020
ACATTTGAGG	GTCTATTAAT	ATTTTCTTTA	TTAACAATTT	CTAAAGCAAT	TGACAATGCT	52080
TTTAGAGAAT	CCTTATCTAA	AAGGTAAC TA	AACATTGAAA	TTTTAAAATT	ATTGTCAATC	52140
TTAAAATCAA	ACATAATGTT	TTTTATCTCA	ATATTCCCAA	GATCCATATC	AATTAAGGCC	52200
TTAGCAGAAG	CCTCCCTAAC	TTTAAGAGAT	GGATCGCTTT	TAAGCTTATA	AATCAAAATA	52260

TCCTTTGCAG	AAGAGTCCCT	ATGTCCTTTG	ATTGCATTAA	TAGCTTTAAA	CCTAATATTA	52320
TCATCAGAAT	CTCTTAAAAA	TCCTTGTA	ATCTCTTTAG	ACTTTAAAGA	AGGATCTTTG	52380
GACAAAGAAG	CAATAATAGC	TAATTTAACA	TTTAAATTAT	TGTTATTACT	CTGAAGATAC	52440
AAATCAGCAT	TTTCAGTTAC	TTTATCTGAA	GCAAGATATG	ACAACGCTTC	GATTGCAGCA	52500
GCCTTAATTG	ATGGGCCCTC	GTAATTATCT	AGCGAAATTT	CATAAATTCT	ATCCTGATAA	52560
TCAACAGCGG	ACATTTTTTC	AAGAGCAATA	AGTATTTCTC	TTCTAGCCCC	ATCATTTTCCA	52620
GAATATTTTT	CAAAAACCTC	CATCATGTTT	TTAGAATACT	CAAGAGAATT	AAGCTCTCCT	52680
AAATAATAAG	CTGCAATAGA	TACCACATTG	CCCTCTTTAT	TTTCAAGAAT	GTCAATAAGA	52740
GTTTTTTTTA	ATTTTTCTTT	ATCATCAAAC	TCCTTAAGAT	ACGAAATTGC	CAAGCCAAAT	52800
AAAGCGTTTG	AATATCTTTT	ACTCTCATAA	TTTTCAAGAA	TATAATTTGC	TGTATCAATG	52860
CCCCCGAAT	ACTTAAGAGA	AATAAACAAAT	TCAAGTATTT	CCCTTTTAAG	CTCAGCATTAA	52920
AAAGTTTTCT	CAAGTCTTTT	TTTAAGAGAA	AAATTATATT	GACTATCGCT	TGATTTTTTA	52980
AGAGCTTTTA	TAATGCTTGT	CACTTGACTA	TCAAGCCCAT	AAAGAATTGT	ATCGTTAACA	53040
TACTTACCAT	CTAAACCAAC	ATTAGAAAAA	TTCTCTCCCT	TAGAAGAATT	TTCTCTCTCA	53100
ACAGGCTTAT	TTTCTGTAAT	TTCGGGCAAC	AAAGGCGGAC	TAGGAAGAGC	TGGAGAATTA	53160
ACATTTTGAG	CATACACATT	AAAAATAAGT	AAAAAAAATA	AAAAATAAAA	GTATTTTATA	53220
AAGCATCCCT	TCTAATATAT	CTAAAAAGCT	TATTTATTCC	TAAAACAGAA	TAACAAATAA	53280
ACAAAACAAA	AATACTAACA	ATTCCAGCTG	CCATTAAAAA	ATAAAGATTT	TTAAAACTAA	53340
ACCCACATC	CCACTGAAAC	TTTTCAAAAA	AGAAATAAAT	TGCATATAAA	GGAAAAAGTG	53400
TAATAATTGA	CTTTAAAAAG	ACAAATAAAA	TTTCAATTAA	ATCAATTTTA	ACTCCTCTTT	53460
TCAATATTAT	AAAATAAAAA	ACAATTACAC	AAATCATAAA	AGAAATAGAT	TGAGCTAATG	53520
CTAAAGCGTT	CAAACCATAA	TAATTAATAC	CAAAAACAGA	TATTGCAATA	TCAAGAATAG	53580
AAAATAAAAC	ACTCAAATAA	AACGGTGTTT	TTGCATCACG	AATAGAAAAA	TAATATTTTT	53640
GGAAAAAACC	AAACATTGAA	TAAAAAAGCA	GACCTAAAAG	AAAACATTTT	AAAACACTCG	53700
CTGTTTTTTG	AGTATCATAA	ATAGAAAACT	TGCCTCCCAT	AAGAAATAAA	TTTAAATAT	53760
AATCAGACCA	AATAAACATT	AAAAAAGACA	CTGGAATAAA	AATTAACAAT	AAAATTTTAA	53820
TTCCATCTAC	TAAAAGGGCA	TTTAATTTTA	TATTATTCCC	CAAAACAGCA	TGCTCTGCCA	53880
TTTTGGGGAA	AATCACTGTT	GCAATAGAAA	TATAAAAAAT	TCCTACAGGA	AGCTGATAAT	53940
AAACTACAGC	ATTACTAAGG	ATAGAAACAC	TTCTATCTC	AAGAGTAGAT	GCTAATGCAA	54000
ATGAAATCTG	CTGAGTAATA	ATTGAAATGG	GAAAATCCAA	GAATCATACG	AAGCCATCTG	54060

GTAAAAAATT	TAAAACCTTT	TCTCTGAAAT	AAAATGTTGG	CTTCCAGGCA	AAACCAATCA	54120
TAAGGCAATT	TGCAAACGGA	ATTAAAAATT	GTAAAAACCC	CCCAAAAATT	ACGCCAATAA	54180
CAGCACTATA	TATTCCAAAA	CGACCATAAA	ATAAGAATAT	GCTCAATATT	ATTCCAAAAAG	54240
AAAGCATAAT	GGGCGAAAAC	GAAGGAATGA	AAAAAATTTT	ATATGAATTT	AGAACAGACA	54300
CGAAGATTGA	TGATAGGCTT	ATTAGTAAAA	TATATAATAC	CAAATAACCA	AATACAGAAC	54360
TTGCAAAAAAT	TAAGTTTTCT	CCCCTATAAT	AAGATATAAA	ATACATAATA	GGCTTTGCAA	54420
AAATAATCAT	AACTAAAACA	ATTAACCCAA	TAGAAATAAT	GTAAAGGTT	ATGACAGTTC	54480
TGAAAAAAGA	AACAGCTTTT	TCGTGCGATT	TGTTTTTTTC	ATGTGTAAAT	TCAGGCAAAA	54540
AAGCCGAGGT	CATCGCGCCC	TCTGAAAGAA	TTTTCGCGAA	ATTATTAGGA	ATATTGAAAA	54600
CATAGTTAAA	AATATCAGCA	TCAAGATTTG	CACCAAAATA	ATAAGAGAAA	ATCTTTATCT	54660
TTACAAAGCC	CATTATTCTT	GAAAAAAAAG	TGGAAATCAT	GACCAAAATT	GTAGAAACAA	54720
CATATTTATT	CATCGAAATT	TTCCTCTTCA	TACTTTTTTA	AATATGAAGT	TCTAAAATTT	54780
AAAAAATCAT	CATTTAGAAT	TGCGGCTCTG	ATCTTTGAAA	TCAATCGAAA	CATATAGTGG	54840
ATATTATGTT	CACTTGCCAA	AACTATTCCA	AAAAGCTCTT	TCGATTTTAT	TAAATGTCTT	54900
AAATATCCTC	TTGAATACCT	TTTACATAAA	GTACAGATGC	AATTTTTCTC	TACCTTAGAA	54960
GTATCATCCT	TATACTCCTT	TCTACCAATG	CACATAATCC	CATTATCTGT	CAAAAGAGAC	55020
CCATGCCTAG	TAATTCTTGC	GGGATTAAAG	CAATCAAAAA	TATCAATGCC	ATAATATATG	55080
GCATTAAGTA	TGTAATGGGG	AGTGCCAATA	CCCATTACAT	ACCTTGGTTT	TTCTTTTGGT	55140
ATCAACAAAA	AACTATATTC	AAGGATTCTT	AAATATTCTT	CCCTTGGTTC	TCCAACAGAA	55200
ATGCCTCCAA	TGGCAATACC	TGGGCTGTCT	AATTCCAATA	TATCATTGAT	ACTTCTTTTC	55260
CTTAAATCTT	TAAAAAAATT	TCCTTGAGTT	ATTAAAAATA	AAAGCCCGTT	GTATCCCTCT	55320
TTTCTGTTTT	TAGAAGATTT	GAACGTGCTG	CTAGCCCAAT	TGGTTGTAAT	ATTTGTATAT	55380
AAATTGGCTT	CATTATAATC	AATCCCATAA	GAAGTGCAAA	TGTCAAGTGG	CATAATAATA	55440
TCACTGCCAA	AAATTTCTTG	CATAGCAAAT	ATTCCCTCGG	AAGTAAAATA	ATGGTACGAT	55500
CCATCTATAT	GAGATTTAAA	ATGCACACCT	TTTAGATCAA	TTTTTCTCAG	ATCAGAAAAA	55560
GAAAACACCC	GAAATCCGCC	CGAATCGGTT	AAAAAATTTT	TATTCCAAAT	TGTAAAATTA	55620
TGAAGACCAA	CATATTTTTT	AACAGTTTTA	ATTCCCAGCC	TTAAATATAA	ATGATAAGTA	55680
TTTGCAAGCA	TCAAATTACA	TTCTAACTTC	TCAAGAACAG	CATGTTTTAA	CCCTTTCATT	55740
GGCCCCAAAG	TACCAACTGG	CATAAAACAA	GGAATATCTA	CTCTACCATG	AGGAAGATTT	55800

AAAAATCCAA	CCCTTGCATT	AAAATGCTTA	TCATTCTTGA	TTACACTAAA	CATATAAATA	55860
TCCCAAATAA	TTATATATAT	TATTTCCCTAA	CAATCCTATA	AACAAAAAAA	TTGATAATCA	55920
AAAAAATAT	AGTTGTAAAA	GAACTTGCTA	TATATATATG	CAAATAGCCT	AGATCCGCTA	55980
AAAAATTAAA	AATTACAATA	GAAATAACAT	AAACAACAGC	AAAAGCAATG	CTATTTAATA	56040
GGCTGAGTAT	AAAAATATTT	TTTTTAAGAG	CAAGAGCAAT	AAACCCAACA	GTAAAGCTAA	56100
GAAGAATCAA	TCTAAATGAA	AAGAATATCC	TATTTAATAA	ATCAAAAAAT	GCATCAGAAT	56160
AATTTAAACG	TTCAGCTTTG	AGAAAACTTA	TCCAATTAAT	AAGTTTAGTA	AAATTTAACG	56220
CCTTTGATGA	GAGCATCACA	GTTCTTATGT	AATCGGGCGC	CAGCTTAATA	ATTCTGTGCC	56280
CATCAAGAAC	ATCGTAGGCG	TTCTCCTTAA	TTTTTTTACC	AACCTTAACA	AACTCTCTAA	56340
TACCATAAAG	CCTCCATTTA	TTATCTTTCC	ATTCGGCTTT	ATTTATATCG	TACCTTGTTT	56400
GAAACTCATC	TTTATTGTCT	TTAATTATAA	TCATCAAGTT	AGCAAAAGTA	TTCTCATCAA	56460
TATCATAAGA	TTTGATATTA	TAAATTTCTC	TAGCAAAATC	CCTTATTATT	ATAGTTTTAT	56520
CCCCAGATCT	ACTGTCGCCA	ATGCTATTCT	TAATAAGAAC	ATCTCTTCTT	GCTATAGTAT	56580
CTATTACCAA	ATAATTATCA	AAAAAGAAAA	GAACAACTGA	AATAAATATA	CTAATTAAAA	56640
TAATTGGTTT	TAATATCCTG	GTAAGTGGA	CTCCACAAC	AAAAAGACCT	ATTATTTTAT	56700
TTCTCATAGA	AAGATTGCCA	ATAAGATTCT	AAATAGCAAA	AAGAAAAGAT	AAAGCCACCC	56760
CATCTGAGAA	TGCCTTTGGC	AAATATAAAT	AATAAATATA	AAGAATATCC	TTAAGGCCAA	56820
TATTCTTTTC	AAGATAGTTA	AGAAGATTAA	CAAACAAATC	ACCAAGCATA	ATTAAATCA	56880
TGAAAAGCAG	GTTTCATGGAC	AAAAAAGTAA	GAATGATGCT	TTTTATAAAA	AGCTTATCTA	56940
TTTTTCATTT	TTTAACAATC	TCAAAAAGAG	AATTGCTCCT	GCAATAATTA	AAATTAAATT	57000
AGGCAAAATA	GTAACAATAA	TAGGACTTGG	TGCATACTGC	ACAGTATAAA	CTTTTCCACC	57060
AATAAACATT	ACCCAATAAA	AAACACAAAC	AATAATTGAA	ATTACAAGTT	CAAGAATAAT	57120
GGAATATTTT	CTATTAGAAT	ACATTCCCAT	TGAAAAAGCT	AAAAAAATAA	AAAATAAAAC	57180
TGAAAGTGGT	AAACTAATTT	TTTGATAAAA	TTCAAGATTA	AACAGAGCCA	AATTTTGCTT	57240
CATGCTTCTA	TCTTGATAAG	GTTTGAAATT	TAAATTTAAA	TTATACATAT	AGTTTAAATT	57300
TTCAAAAACA	TAAGATTCAT	CCACATAATA	GTTTTGATTG	TATAAATAAT	TTAAATAAAG	57360
ATTTGAAAAA	TTTAACTTA	AAAAGTCTCC	TTCTAGATTA	TTTTTTATAT	TTGAATCTGC	57420
AATTAAATTA	TTTTGCTTTT	TAATTAATTT	TATAACATCT	CTCATGCTCA	TTTGTGAAGG	57480
AGTTACATAA	TTTAATAAAA	AACTATCACT	AAATGTAACC	TGATCGATTG	AATATTTTAT	57540
CTTATCTGCA	TAAAAATAAT	CATAAAATCC	ACTCTCACTG	TCTGTTAAGG	CAATAGATAG	57600

AACATCATTTT	AAAATAAAAT	ACACTTGAAA	ATTTTCTTTT	CTAATATCAA	GATTTTTTGC	57660
CATAAATATT	CTATCAAAAC	CCTTAAGCCC	AGTGTTATCA	AAAAAAGTTA	CATTTTTTATA	57720
ACCATTTTCC	GATTTCTCAC	CAGAAACAAA	AATCAAATCT	CCATATTGTT	TGCTTGAATA	57780
AGGCTTTAAT	ACCAAATGGG	GAACTTCTTC	TTTTATTTCA	TTAAAAATTT	TTAATCTGCC	57840
AATAGATCCA	AGTGGAAGTA	AAATATCATT	GGATATAAAA	GATACAAAAG	CAATAACTAT	57900
TCCCAATTTA	AAAAATGGGA	CAAGTAAATC	AAAAATTGAT	ATGCCAATTG	AACGAAAAGC	57960
TAAAATTTCA	TTGTGAAGCT	TGAATTTATG	AATAGTAAGA	ATTACTGAAA	TCAAAGAAGC	58020
AAAAGGGGGA	GAAAGCGCAA	TAACCATAGG	AAGAGAATAT	ATAATAAAAA	TAAAAGCCTT	58080
AAAAAAGGGA	ACATAATTTT	GAAGAAGTAT	TCTCATAAAG	AATAAAATTT	GATTTATAAA	58140
AAATACGAAA	AAGAAAAATA	AAAACGTAAT	TAAAAAATAT	TTAAAAAATT	CAGCAATTAT	58200
GTAAGACTCA	TAACTGTTTT	TTAATATTTT	CATCTACAAA	AACCAAACCG	TTATTAATAG	58260
TGCCCAGTAT	GACATAATTT	TCAAATCTAG	AAACTTTTAT	TAAATTCAAA	TTAAGAAGCC	58320
CATTATTGGG	TCCAAAATAA	TCCCAGTTGT	CATTTTCAGA	ATCATAAATC	AATAACCCAT	58380
GATCAAAGGT	TGCAAACAAT	AGCTTTTTTAT	CTTTAATCTC	CATATCCATA	AAATAATTAA	58440
CATCAATATT	ATTGGCAATA	ACGTGCTTTT	TGTAACATTT	TTTATTTAAA	TTAATTCAA	58500
AAAGACCCCC	ACCATATGTT	CCAACAAAAT	AACTATCTTT	ATATTCTTTT	ATAAAATTAA	58560
TATTTTTTTC	ATTATCATTT	TTGCTAAAAA	AATCCAAATG	TTCAATCTTT	TTCAAATTAT	58620
CGACATTAAC	ACTATAAATA	GCCTTGTCAA	CTGTTCCAAC	TAATAATAAA	TTTTTTTAAAC	58680
TATCAAAGCA	GAGTGAAGAA	ATTTTATTAG	ATCCAAGCGG	TATATTTTTT	CAATTTTTTA	58740
AATCATAAAA	CCATAATCCA	GAATTTAGAG	TGCCAACAAA	TATTCCATTT	TTAACAGCAA	58800
GCAAAACTTG	TACATTGCTA	AAATCAGCAT	TACCGGGAAC	ATTTATTTGC	TTTAAATCCC	58860
CATCAACATC	ATCTATATAA	TAAACAACAT	TTTTACCACC	AATATAAATT	GTTCCATTAT	58920
AATCCGCAAA	ACCCCTAATG	CCATTTAAAA	AAATGCTTTT	TTTATCCTTA	AGATAGACTC	58980
TACAATCATT	TTTTTTAATA	TTATATCTTA	AAAGCCCTCC	CAATATATTA	GTTACAAATA	59040
TATTGTCATT	AAAGACAAAT	GTATCAAAAA	CGCTGTTGTC	AAGAAATCCT	AAAGACTCTA	59100
AGTTTAAATG	ACTTACTCCA	AGTTTATTAC	TTAAAAAGCC	ATAAATCTCT	CTATCATAAC	59160
CTGAGATAGA	AAATTCATTA	ATCTCTTTAA	ATGCAGAAAT	TGCATCAGAT	TTTTCTTTAA	59220
CAAGATATTT	TAATTCAGCT	AATCTTAAAC	TAGCATGAGA	ATATTTATAG	TCTTTTAAAA	59280
ATAGATCAAA	ATTATATTCG	GAAAGATCAT	AAAAACCATT	CTCATAATTT	ACATATCCAA	59340

AAAACAAATT	CGCTTTAGCT	AGCAATTCTC	TGCTGTGCTG	ATTCTCATT	GTTACTATTT	59400
TATTCAAATA	ATAATTTGTC	AAGCCAATAT	TTTTTTTAGC	ATAACTTTCC	CTGGCTTTTT	59460
TAAAATAAAA	ATTATTTTCT	TTTAAAAAAG	CATCACTAAG	AAAAAAAGAT	TCATTATCTT	59520
TTAATCCTAT	CAAATAGTCT	CTTTTAAATT	TACCCTCATT	ACTCCCAAGA	ACAACATTAT	59580
TATCATCAAT	AATGACTGCT	TCTTTTTTCT	TCTCTACAAT	ATCACTAATA	TGAGAATCTT	59640
GAATAGATCT	ATCTGTAGTA	AGGCAGGAAA	AAAACAAGAA	AAAACATATA	AGACAACCTT	59700
TAAATAAATT	ATTCAAAACA	AATTTTCATAT	TATTTTAATA	ATCTTTATCT	CTAACAATTT	59760
CAAGATCAAT	TTTTCCAAC	TTGTCTATAT	CAATTATTTT	GACCTTAATT	CGCTGACCTT	59820
CTTCTAATTT	TGGGGGTCGA	ACTAACCCCG	CATTACCTCT	TATATTCTCT	CCACCGCCAC	59880
CAAATCTAGA	ATATCTATTA	CTATTCCCAA	ATCTTCCAGA	ACCATACTTA	CTGTCTCTGG	59940
GTTTCAAACG	AGTACTTAAA	AATCCTTCCT	TTGCAGGAGT	AAGTTCAATA	AAAGCCCCAA	60000
AGCTATTAAT	CTTTTTGACA	GTTCCCTCAT	AAATTCGCC	TACCTTTGGC	TCTCTTACAA	60060
TACTCTCTAT	TCTTTCTTTA	GCTTTTTGCA	TCTTAAAATC	ATCATCCCCG	AAAAGAATGA	60120
TTTTTCCATT	CTGCTCAATT	TGAACCTTAA	CTTCAAATTC	ATCTGTTATA	GCCTTAACAG	60180
TTTTTCCAGT	AGATCCTATC	ACAAGAGATA	TCTTGTC AAT	GTCAATTTGA	AGTTGAACAA	60240
TTTTTAGGAGC	ATACTTAGAT	ATACCAACTC	TTGAATTAGA	AATTACAGTA	TTCATAATAG	60300
ATAATATATG	TATTCTACCT	ATTCTTGCTT	GCTCAAGAGC	ATCTCTCATT	AAATCTTTAG	60360
TAACATTTTC	AATCTTAATA	TCCATTTGAA	ATCCAGTAAT	TCCATTTTTT	GTACCGGCCA	60420
CTTTAAAGTC	CATATCACCT	AGATGATCTT	CTTCTCCAAG	AATATCACTT	AAAACACAT	60480
ATTTATCCCC	TTCGCTAATA	AGCCCCATGG	CTATCCCCGC	AACCTGCCCT	TTAACAGGAA	60540
CCCCTGCTGA	CATTAAAGAC	ATGCTCCCAG	CACAAACAGT	AGCCATTGAA	GAAGATCCGT	60600
TAGACTCTAA	AACCTCAGAA	ACTACCCTAA	TGGTATAAGG	AAAATCATTT	TTTCCAGGAA	60660
CCATTGATTC	TAAAGCTCTT	TGAGCTAAAT	GACCATGGCC	AATCTCGCGC	CTGCCAGTCA	60720
TTAGTCTACC	GGTCTCACCA	ACTGAAAATG	GGGGAAAATT	GTAGTGGAGC	ATAAAATTAA	60780
GGCGTTTATC	GCCATCAATA	TCATCCATTA	TTTGTTTCATC	AATGCTTGTA	CCAAGAGTAG	60840
TTACCGCTAA	AGCTTGCGTC	TCTCCCCTTG	TAAAAAGCGC	AGATCCATGC	GTTCTACTTA	60900
AAATATCAAC	TTCTGAGATA	ATATCTCTTA	TCTCATTAGG	AGTTCTGCCA	TCTGTTCTAA	60960
TATTATCGTT	AAGAATAGAG	CTTCTAACAA	TCTCCTTCTC	AAAATCATCA	AAAGCCTTAT	61020
GAAAAAGAGA	TTCATTGCTA	TCAGTCAATT	TCTCAAGAGA	AGAAAAGTAC	TCATAAGATT	61080
TATTTGCGAG	CAAAGTTATG	GCTTTATCTC	TATTAAGCTT	TCCCTTAACA	AAACAAGCTT	61140

CTTTAAGATC	AGCATAAACA	AAATCCCTAA	GCTCATCTTT	AAATTCAAAT	ATTTTTTCTT	61200
CAAAAGCTAA	AGGAAGTTTT	TCCTTCTTGC	CTACAATATC	TAAAAATTCT	TTTTGAGCAT	61260
TACAAATTTG	CTTAATATAT	TCATGAGCAC	CATCTATTGC	TGAGAGCAAA	ATATCCTCAC	61320
CAACCTCATT	AGCACCACCT	TCTACCATAG	TAATTCCATT	TAAACTTCCG	GCAACAACAA	61380
TATCAAGATC	AGAATCATGA	ATCTCTTCAA	ACGAAGGGTT	TACTATAAAC	TTACCATTCA	61440
AATAAACCAT	TCTAACAGCT	GCAATTGGAC	CATTAAACGG	AATATCTGAC	AAAAAACTG	61500
CCGTAAAAGC	AGCATTTCATT	CCAACAATAT	CAGGAGGATT	AAGCTGATCT	GTAGCTAAAG	61560
TTGTAGGAAT	TACTTGAATT	TCTCGACCAA	ATCTTTTATC	AAAAAGAGGT	CTCATCGGCC	61620
TGTCTATTAG	TCTGGAAACA	AGTATTTCTT	TATCCTTTGG	CTTTCCTTCT	CTTTTGATAA	61680
ATCCTCCCGG	AATTTTACCG	GCTGCATAAT	ATTTCTCATT	ATATTCAACA	GAAAGCGGAA	61740
CAAAATCTAA	ATCTTCTCTC	ACGTTACTCG	AGCAACAAAC	AGTTGCAAGA	ACCGAAGATC	61800
CACCATAAGT	TGCAAGAACC	GATCCATTAG	CCTGTTTAGC	CATAAATCCG	GTCTCAAACA	61860
CTAACTCGTC	TCTGCCTATT	TTCAACTTTA	ATATTTTCCT	CAAAATTCAA	CCTCTTTTTA	61920
TTTTCTAAGA	CCAAGTTTAG	ATATCAACAT	CCTATAAGCT	TCTAAATCTT	TTTTCTGGTA	61980
ATACCGCAAT	AAACTTCGCC	TTTGCCCTAC	TAACTTTAAC	AAGCCTCTTT	TTGAACTATG	62040
ATCTTTTTTA	TTTATCTTTA	AATGTTTCAGT	TAAATACTTT	ATTCTACCTG	TAATAAGTGC	62100
AATCTGAACC	CCAACAGAAC	CAGTATCACT	TTCATTTTTT	CCAAATTCAG	AAACTATTTT	62160
TTGCTTTTGC	TTTTTATCTA	TCATAAAGCA	ACTCCTATAC	CATTATAGCA	AAGCTCTAAC	62220
AAACCTCTTG	CCATAATTTA	AAATAACTAC	TACGATAGAT	TATAATATTT	TTTCTTAAAA	62280
ATAACAAAAG	CAATTTATCC	TTTTCGGGTT	ATTTTTAATA	ATAAATTATT	AAATTGTTTA	62340
AAAAAACAAA	TATAAAATTA	AAATGTCATA	ATATATTTAT	AATTAAAATA	TAATGACATA	62400
TTTATATTTA	TTCAAACCCA	CTCCTTGAAT	TACTGCTAAT	ATTTTCTCTT	CTCTGGATTT	62460
TAAAATTTTA	AATTCATTAA	TATTGATTTC	AATTTCAAAA	TAAACACCAT	TTTAAACAAG	62520
ATTTATCTTA	TTAGAATCAA	TGTAAACCTT	TTCAAACTT	TTTAAAGATT	CTAGACTAAT	62580
TAAGGAAGCT	TTACTCAAAT	TTTCACACAA	TGTGGAATCT	TTTAATCTAA	ACATACCTAC	62640
TTTAGTCCTT	TTTAAATTAC	TAACATACGC	GCAAGAATTT	AGAGAATATG	CCAAATCTCT	62700
TGCAATACTC	CTAATATAAG	TACCTTTTGA	ACAGCTAATT	TTCAAACTAA	GCAAAGAAGA	62760
ACTAAAATCA	TAACTTAATC	TTTGAATATT	ATAAACAGTG	ACTTTTCGTT	TTTTAATTTT	62820
AAAAAACTTT	CCATTCAAAG	CAAGTTTATA	GGCTCTGCTG	CCATCAATAT	GAACAGAAGA	62880

AAATCTAGGA	GGACTTTGAT	AAATCTCTCC	TACAAAATCT	TTAAGCTTTA	AATCTATATC	62940
CTCTACATTA	GGAATATAAT	CTGTTTTACT	AACATTCTCT	CCATTTCGGAT	CAAGGGTATC	63000
TGTTTCTAAT	CCAAATCTGA	ATTCTGCTAC	ATACTCTTTA	TCTAAAGAAG	TAAAAATAACC	63060
TGAAAGCTTT	GTGTATTTTC	CCACAAGACA	AACCAAAATT	CCACTTGCAA	ATTTATCAAG	63120
TGTGCCAGCA	TGCCCAACAC	GATTTGTATT	AAAATATTTT	TTTATAGGGA	AAAGAGTTTC	63180
AAAAGAAGTT	TTACCTTGTT	CTTTATTAAT	TAAAAGGAAT	CCATTTTCCA	AATTTAATTC	63240
TCTCTTGTAG	TATTTAATCC	TTCAATTAAC	TTATTAACAT	AAAATGATTT	GGAAAGAGAA	63300
TCATCCTTAA	CAAATAATAA	TTTGGGAGTG	CTTCTAACTT	TAATTGCTTT	AATAATTTGA	63360
CTTTGAATAA	ATCCCTTAGC	ATTATTTTAA	GCTTTAAGTG	CATTGTCCAA	AGAAGCACCT	63420
TCCTTAATAG	AGCCCATAAA	CACTTTAGCA	TTTATTAAAT	CTTTTGAAAA	TTCTACTTTA	63480
ACCACGGTTA	AAAATGAATG	AATTCTGGGA	TCTTTAATCC	CCCCACTTAC	TATTAAATTG	63540
CCGATTTCTT	GAGCAATAAA	ACTTTCAAGT	TTAAACTTTT	TAATATTCTT	ATACATAAAC	63600
ACATATAAAT	AAAACAATAC	TAAGTTTTTA	AAGATTTTTT	AACCTTTTTT	ACCTCAAATG	63660
CTTCAATTAT	ATCTCCTTCT	TTAATATTAG	CATAATTATC	AATCATAATA	CCACACTCAT	63720
ATTGCTCAGC	AACTTCTTTA	ACATCATCTT	TAAATCGCTT	TAAAGATGAA	ATTTTGCCGG	63780
AATGAATCTG	TAAACCATCT	CTCATTACAT	TAGTAATCGC	ATCTCGCTTT	ATTAGCCCCC	63840
GAGAAACATA	ACAACCGGCT	ATTACCCCTA	TTTTAGGAAC	ATTTATTACA	GCTCTCACTT	63900
CAGCAAAGCC	AATAAACTGC	TGCTCAACAT	CTGGCTCAAG	CATTCTTTCA	AGAACTGACC	63960
TAACATCATT	TATAGCATCA	TAAATAACAT	TGTACTTTCT	AATCTCAACT	TTTTCTTGAT	64020
CTGCTAGTAC	CTGAGCTTTT	GCAGTAGGCC	TTACATGAAA	TCCAATAACA	ATAGCATCGC	64080
TTGCTGAAGC	AAAGCTAATA	TCTGTTTCGG	TTATTACCCC	TGCTGATGAA	TGCACAACCTC	64140
TTACTCGAAC	CTCATCGTTT	GTTAATTTTT	CAAGAGAATT	CTTTAAAGCT	TCCACTGAGC	64200
CTTGAACATC	TGCTTTTAAA	ATTATTTTAA	GCTCTTTAAG	CGCTCCTTCT	TTAATTGAAT	64260
CATAAAGATT	CAACATAGTA	ACTTTCTTTA	CATTTTGGGA	AGATTCATAT	TTTTTAAGAT	64320
CTTGTCTTTT	AGAACTGATC	AATTTTGCTT	CTTTTTCAGT	TTTAGTTACT	TGAAAAGGAT	64380
CCCCGGCTTG	AGGCATTGAA	GAAAATCCTA	AAACACTAAT	GGCTTTAGCG	GGTCCAACGC	64440
TCTTAACAGA	AACACCCTTT	TCGCTAATTA	ATGCCTTAAC	TTTACCATAG	CACGCTCCAC	64500
CCACAAAAGA	ATCTCCCACA	TAAAGCGTTC	CATCCTCAAT	AATAACAGAA	CAAACTATTC	64560
CGCGCCCCAA	ATCAATCTTG	GCATCAAGCA	CTTTTCCAAT	AGCTCTTTTG	GATGGATTTG	64620
CCTTTAACAA	CATCATATCT	GACTGTAAAA	GAATCATATC	AAGTAGTTCA	GAAATTCCTA	64680

TATTTTTAAG	AGCAGAAATC	ATCACAAAAA	TAGTATCTCC	CCCCCAATCC	TCAGATACTA	64740
AACCGTATTC	TGAAAGCTGG	TGTTTAATCT	TATCGGGATT	TGAATCTGGT	AAATCAATCT	64800
TATTTATAGC	AACAATAATT	GGAACATTTG	CCTCTTTTGC	ATGATTGATA	GCCTCAATGG	64860
TTTGGGGCAT	AACACCATCA	ATTGCTGACA	CAACAAGAAC	AACAATATCT	GTAACCTGAG	64920
CCCCACGACT	TCTCATCATA	GTAAAAGCTT	CATGACCAGG	AGTATCTAAA	AATGTTATTT	64980
CTCGATCATT	ATAAACAATA	GTATAAGCTC	CAATATGCTG	AGTAATACCA	CCGGAATCTG	65040
TTTGATTTAT	ATCTATATTT	TGAAGCACAG	AAAGTAGTTT	GGTTTTGCCA	TGATCAACAT	65100
GACCCATTAT	TGTAATAACA	GGAGGCTTTT	CAACTCTTTT	GCTTTGATCT	TCCACTTCTT	65160
CTTCTATAAC	CGTTTCATCA	TAAATAGAGA	CAACATTAAC	TTTTGAACCA	TATTCTTCAA	65220
CTAAAATAGT	TGCAGTATCA	GAATCTATCT	TTTCATTAAT	AGTAACCATT	ACGCCCCAAG	65280
CCATTAATTT	AGCAATCAAA	TCAGAAGATT	TTAAATTCAT	CTTCTTGCA	AGATCAGAAA	65340
CAGTAATGCT	ACCCATAATG	TCAATTGACT	TTGGAATAGG	GTTGGCTAAA	TTTTCTCTCT	65400
TCTTTTTCTG	AAGTTGTTCA	AAAACTTTTT	GTTCAATTGT	TTTGCTCTCA	GTTTCTGCTT	65460
TTTTTCTCTT	ATAGCTTTTT	TGACTCTCTT	GTTGCTGTTT	TTTTTCTCG	CCAAGCTTAC	65520
GATTTAACTC	TTTACTATTC	TCAGAATCCG	CTGCAGGTGT	GCTGCTAACA	ATAGCGGGAA	65580
CTTTAGTTTT	TATAAGTCTT	CTAAAAGACA	TAGAAGTAGT	AGTATATTTA	TTTTGAGAAT	65640
TATTTTTGGC	AACATATGTT	TTCTTTACTG	AACCTTGATA	TTGAAAGGAT	AAGCTGTCTC	65700
TGTTTTGTGA	ATATCCACCA	GTTCTGTTAT	CTCTGTTTTG	TGAATATCCA	CCAGTTCTAT	65760
TGCTCTGTGT	TTGTGAATAT	CCACCAGTTC	TATTGTCCCT	GTTTTGTGAA	TATCCACCAG	65820
TTCTATTGTC	CCTGCTTTGT	GAATATCCAC	CAGTTCTGTT	GTCTCTGCCT	TGTGAATATC	65880
CACCTCTATT	ATCCCTATTT	TGTGAATACC	CACCAGTTCT	GTTGTCTCTG	TTTTGTGAAT	65940
ACCCACCAGT	TCTGTTGTCT	CTGTTTTGTG	AATACCCACC	AGTTCTGTTG	TCTCTGTTTT	66000
GGGAATATCC	ACCAGCTCTA	TTGTCCCTAT	TTTGTGAATA	TCCACCAGCT	CTATTGTCCC	66060
TATTTTGTGA	ATACCCGCCA	GTTCTATTGT	CTCTACTTTG	CGAATATTCA	GCCTTATTGC	66120
TGTTATTATG	CAAATCAACA	AAGCTATTTG	AATCATTTTT	AACGCTTAAA	TCATTATATG	66180
TTACAATTTT	TACTACCTTC	TTTTTCAACT	TAATAATCTT	AACTTTTTTG	CCATCTTCAT	66240
TTTTAATATC	ATCAATATTT	TTGACAAAC	CTACTCCTCC	TCAAACCTCA	AACTAAGCCC	66300
TATTTTACAA	CCCGGACAGG	AGGTCATATT	TTCATTAATA	ACAACACCGC	ATTCAGGACA	66360
AAGAAGCTCT	TCATCTTCTT	CTACCTTTTC	CATAGACTCA	TCATTGTCAT	TAGCAATTAT	66420

TATCATCCCC	TCTTTTAATA	TTTGTGTTAAT	TTCTTCTTGT	TTTTTCATAAC	TTACACCAAG	66480
ATTAAAAAGC	ACTCCCTCAT	CTGCTTGTA	AAAATTGTGA	ATATCATCAA	ACCCCTCTTT	66540
TGATAAATTA	GAAATCACAG	AAGGATCAAG	CAATTTAAGA	TCACTTATTT	TACTAATCTC	66600
TTCAAATTGC	TCCTCCTCAA	CAACATCTTG	CATAACTTTA	TCAAACATTT	CGAGTGTTTC	66660
TTGCTTAAAC	TCCGAATTAG	CCTTCATTTT	TGCAAATTGA	CTGCTAGTTT	TAACATCAAT	66720
AGCCCAGTCA	AGAAGTCTAT	TAGCAAGTCT	AACATTTTGA	CCCATTTTAC	CTATAGCAAG	66780
AGAAAGCTGG	TCATCACTAA	CAACCACTAA	AGCTTTATGT	AAATCCTCGT	CAAGAATATA	66840
AACATGTTCT	ATCTTTGAAG	GAGTCAAAGA	ATCCTTTATA	AATTCCTTAA	TATCTTTACT	66900
ATAGGGAATA	ATATCAATTT	TTTCTCCTTC	AAGTTCCTTA	ATTATAGATT	GAATTCTGAC	66960
TCCTTTTTGT	CCTATACAAG	GACCAACAGG	ATCAATCTCT	TCTTTTTTCAG	AATAAACAGC	67020
GACTTTGATT	CTGTAACCAG	GATCGCGAAC	TATTTTATGA	ATCTTAATAA	TACCTTCTTC	67080
AATTTCTGGA	ATTTCAAGCG	CTAAAAGCTC	TTCAATAAAC	TTTGGATGGG	TCCTAGAAAG	67140
AATAACTTCA	ATACCATTTT	TACCCTTTTT	GACATTATAA	ACTAAAACCTC	TAATCTTATC	67200
ATTAAGGTTA	TAAACTTCTC	TTGGCGATTG	ATATTTCTTG	GGAATTATAC	CATCCGTATT	67260
ACCAAGATTA	ACATAAAGAT	CACCATTCTT	ATTTTGTTGA	ACGTACCCAA	TAACAACCTT	67320
ATTCAACTTG	CTTTTAAATT	CTGATAAAAT	CTCATTATCC	TCAATTCCTT	GCAGGTCATT	67380
TTTGGTTCTT	TGTTTTGCAA	CCTGAATAGA	AAGCCTATCA	AAAACCTTGG	GATTAATTTT	67440
AATGTAAGCA	TAATCACCTT	CTACAATATT	TTCCCTTGAG	ATATCTTTTT	CTAATATTTT	67500
AAGCAAAGAA	TCTTTTACCT	CTTTTACAAT	TTTCTTTTTT	GCATAAACAG	ACAAATCTCC	67560
CGTATCATCA	TCAAACCTTA	TAAAAGCATT	CTCATTGCTT	CCAAAATACT	TCTTATAAGC	67620
TATTAATACT	GATTCTTTAA	TTGTTTTTCT	AATAGAATCT	ATACTCATGC	CACGATCATT	67680
TGCAATATTT	ACAATCATAT	GCCCCGTGCC	CTTTATCATC	CAAACCTCCT	CCTTAAACTA	67740
ATCTAGCCTT	TTTAACATCA	CTATAAAAAA	CATTTACTTC	TTTGCTATCT	GTTTTAAAAA	67800
TAAAACTTTT	TGGCTTTGAC	TCTAATATAA	AACCGTCTTC	AAATTCATTA	TCCAACATCA	67860
ACTTAATCTT	TTTACCTTCA	AAAATTTTAA	ACTCTCTGTC	ACTTTTTTATT	TTTCTATCTA	67920
TTCTGGAGT	AGAAAGCTCT	AAAGTAAAC	CATATTTAAG	ATTTGCTTCT	AAAATTAATA	67980
AAATCATTTT	ATGCAAATCA	GTCAAAAAAT	CAATATCTAA	GGAAAAATTT	TTACTATAAA	68040
GAACATTTTG	AATTTTTTCCA	TTATTTTTTAT	TTCTAAAGAT	ATTAATTTCT	AATATCTCAA	68100
CATTTAACCG	CCCTGTTAAA	TCTTTTATCA	AATTAAAAAC	TTTATTATTT	TTGTCAAAAT	68160
ACTTAATCAA	CTGTATTCCT	AAAAATAATA	AGGTTCTTTT	AAAGAACCTT	ATTAAATACA	68220

TAAACTAAAC	CTTAAGTCAA	GGTTAACT	TAGCAAAAAT	AATGTCAACA	TAAACCAAA	68280
TTATAAGATT	TGGCCAAAGA	AAGCTTTATC	ACCTTATAAG	CTCCCAATTG	CATAATAAGA	68340
TCTTCACAAA	TGCACATAGA	TGCTCCTGTG	GTAACAATAT	CATCAAGTAA	AACAATCTTT	68400
TTAAACTGAA	AATTTTTATA	TTTTGATCTT	AATTTAATCT	TATTTTCAAG	ATTTTTAAAT	68460
CTAAGATTCC	CTTTCATTAA	CTTCTGGCTT	TTTCATACT	TTCTTGAAAA	AATATTTATA	68520
TAATTAAAAC	CAAAACGGCT	TAACAAAATA	CCAATGTATT	CCATATGATC	AAAACCATAA	68580
AATAATTTTC	TTTTAAAACT	ACAAGGAACA	GTTACTATTT	GATCAAAATC	AATATTATTT	68640
AAACATTGAG	CAATTCCACT	TGCCAAAAAT	CTACCAATTG	ACTTTTGAGC	ATCCCTTTTA	68700
TAAGACAAAA	TAAAGATTT	GTAATGCTCT	TTATATTCAA	AAAAATAAAT	CAAATTCTCA	68760
TCAAATTTAA	TGTTAAAATT	AAAAAGTGAC	TTACATTGGT	CACAAAGAGC	ATTAGAAGAT	68820
ACATACCTTT	TTCCACAAAA	GACACAAAAA	GGCAAAAATA	TACTCTTTAA	AACATTTAAA	68880
TAGCTCATAC	TAAGTGGACT	GAGAAATAAC	CTTTAAACAA	ATTGATTTA	ACAGCTCAAT	68940
TGATTGAAAA	GGAGTAATAT	TATTAATATC	TATGTTAGAA	ATAAAATTTT	TTAACTCTAA	69000
ATACTCATTT	AATTTAATAT	GAATATCAGT	GTCATTTTTC	AAGATCTCTT	TATCATTACC	69060
ATCAGAAGAA	ACATGGGGAA	GAACTCTAA	ACAAGAGTTG	CCCTCTCGGC	CCACCAAAC	69120
TTCTAGAATA	ACATTAGCTC	TATCTATTAC	CCTTAAGGGA	AGTCCTGCTA	TGCGAGCAAC	69180
ATAAATACCA	TAAGAATTAA	GAGATGGCTT	TTCTTCAACT	TCTCTTAAGA	AAACAAGATC	69240
GTTGCCCTGC	TTTTCAATTT	TCATTGAAAG	ATTAATAAAA	GCCTGATGAT	TAATAGACGA	69300
CAATTCATGA	AAATGTGTGG	CAAACAACT	TCTAGCTTTA	ATATACTCTA	AAATATACTC	69360
TATAATAGAA	TAAGCAATAG	CAAGCCCATC	ATTTGTGCTA	GTACCTCTTC	CAACTTCATC	69420
CATAATTATT	AAACTCTTTT	CTGTTGCATT	CCTTAAATG	TTGGCTGTTT	CATTCAATTC	69480
AACTAAAAAA	GTGGATTCCC	CTTGGCAAT	GTTATCACTT	GCTCCAATCC	TGCAAAAAAT	69540
TTTATCTGTA	ATACCTATTA	AAGCTTTAGA	AGCTGGCACA	AAAGAGCCTA	TATGCGCCAT	69600
TAAAGTAATT	AAAGCCACCT	GACGCAAATA	GGTTGATTTA	CCTGCCATAT	TAGGTCCAGT	69660
AATTAAACAA	AAATACTTTT	CTTTATTAAT	TCTTACAAAA	TTTTCAGTAA	AGATTTCACT	69720
ATTTTTAGTG	TAGTGCTCAA	CAACAGGATG	CCGAGACTTT	TCAAGAAGAA	TTTCTTTACC	69780
AGATGTCAAT	ACAGGCCTTT	TATATTCATT	TTTTTTTGCC	AAATAACCAA	AGTTAACAAC	69840
TAAATCAATA	TATGCAAAAA	ATTCTGCAAC	CTTTTAAAGA	ACTTTATTAT	GCATAACAAC	69900
ATTTGATGCT	ATTTTCATCA	AAATTTCCCTG	TTCAAAAGCA	ACCACATTAT	CTTCAGCATT	69960

ATTAATATCC	ACCTCAAGAG	AAATAAGTTT	TTCTGTTTTA	TATCTTTTTG	AAGAATTTAA	70020
AGCTTGGCTT	TCCATAAAAT	GTGGTGGCAC	TTGAGCATAA	TTACTCTTTG	TAACTTCAAA	70080
AAATAACCCC	CTATTATTAG	TTTTTCTAAT	CTTTAGGTTA	TTAATCTTGC	TAAGCAATCT	70140
CTCTGATTCA	AGATATTGAT	CAATATATTT	ATTTGCATTA	ATCTTTAAAT	CTTTTAAGTT	70200
ATCAAGCTTT	AAGTCATAAC	CTCTTTTAAT	AAGTTCATCA	GGTGCACCTG	AAATTGCACT	70260
ATTTATCAAA	AAATAAACTT	TAGAAATACT	ATCCTCTTCA	AATTTATCAA	AATTCCAATA	70320
ATCAAAATTA	TGCTTGTCAA	ATAACTTTTT	TACCGTAAAA	AATACAGAAA	GAGCTTTTTT	70380
AATAAATAAA	AAATCTTTTT	TAATATATCT	TTTCATTTGA	ATCCTAGATA	TTATTCTCTC	70440
AATATCCCAT	ATATTAATAA	AAGTTTCTCT	TAAAGTCACA	GTCAAGCTAA	TATTTTTGCA	70500
AAAAAATTC	ACATGATCTA	GCCTGGTATT	AATCTCAGAA	ATATTTAAAA	TTGGATTTAA	70560
AATAAATTCT	CTTAAAAGTC	TCTTTCCCAT	TGCAGTTTTG	CAATCATTTA	ACACAGAATA	70620
TAATGAATAT	TGAGAAGAAA	AATCATTATT	ATTTTTTACA	AGTTCAAGAT	TAAGTTGAGT	70680
TACGTCATCA	AGAAACATGT	ACGAAGAATC	ATTATTGATA	TCTATTTTAT	CAATATTACT	70740
TAATAAATTT	TTTAAATTAT	TTTTTATATG	ATTTATAATA	AGAAAAATTG	AAATGTAATA	70800
GGGCTTTTCC	TCATCAAATC	CAAGAGAGCT	CAATCCAAGT	ATGTTAAAAT	GCTCCTTTAT	70860
TGTTTTTATT	GCAATATCCT	TATCAAGATG	CCAAGTAGGA	ACTCTGTTAA	TTAAAAATCT	70920
ACTAAGATTA	AGCTTCTCTG	AGTATTCATA	ATAAAAATTT	TCAGAAACTA	TTATCTCTTT	70980
AGGAGAGTAT	TTCTCAAGAT	CCCTTTTAAG	TTTTTCAAAA	AAACCATTC	CATAAAACAT	71040
TATTCCAAGA	CTGGAAGTAG	ATAAATCTAT	ATAAGAAAAC	GAATAATAAT	CTTTATAATC	71100
ACTAATAGCA	ACTAAATAGT	TATTAATATC	ATCATTTAAA	AAATCTTCAT	CAATAATAAC	71160
GCCTGGGGTT	ATTACCTCAA	CAACCTCTCT	TTCTAAAGGC	CCCCCAGAAG	TAGAATTGGA	71220
CGCTTGTTCA	CAAATTGCAA	CCTTTTTATC	AAATAAAATT	AATTCCTTA	TATATTCTTT	71280
ACTGGTATGA	TAAGGAACCC	CACACATTGG	AACATTTTCT	CTTTTGTGCA	ACGTAAATTT	71340
AAGAAGCTTG	CTTACCTCAA	TTGCATCATC	AAAAAACATT	TCATAAAAAC	TTCTACTCT	71400
GAAAAAAGA	ACAGCATCTT	TATATTTTTT	CTTGATATCT	AAATACTGCC	TTATCATTGG	71460
GGTAACATTT	TTTTCCATAT	GCTTCCTAAA	TAATATTGAA	TTACAATTGA	TATTATAAAA	71520
TAAATATAAT	TCAATTAAAA	AGAAAGAATA	TAAAATAATA	AAAAGACCAT	AAAAAAAATA	71580
TTTTACGCAA	TTAAACGCTA	TTTAATTATT	AAAAAGCCTA	ATGTTTTAAA	TTTAATTAAC	71640
TTTAAGGGTT	TTTATTGTCC	TTTTCTAAAA	GATGCTTAAC	AACATCGTTT	GTTATATCAA	71700
CAGTGCTATT	GTGGTAAAGA	ATATATGGAT	TATTTTTTTT	CATAATCAAA	GAAAACCCAT	71760

TAATTTCTGC	AACATATTGA	ATACCACGAA	GTATTTTACT	TAAAGATTCA	CTATTATTAT	71820
TTAAACTATT	AATATTGGCC	AATCTCTGCT	GTCTTAAATT	ATTCTTTGCT	AAACTAGACA	71880
CTCTCTTTAG	CTCATCAACT	TTCAAATTAT	ATTGATTTC	AAAAGATCTT	GCATTATCTA	71940
AATCATTATC	AGCAATCGAT	TTATCATACA	TATGTTTAA	ATTCTTAAGC	TCTAAATTTA	72000
ACGTATTTAT	TTGTTCTTGA	TACCGATTTT	TTATTTGATC	AAGATTAGCT	TTCAATTGAG	72060
GATTTAAAC	TTCAATCACA	ATTCTATCAA	AATCAACAAT	TCCTATCTTA	ATAACATCTA	72120
TCGAAAAAC	ATTAAAAGAT	AATAAAAAA	ACAATGGTAA	AATCAAAAA	AACACAAATT	72180
TCTCCATAGC	ATTAATCAAT	ATCTCATCTC	AATTCCTAAG	AAAAATTTAA	ATCCAGAATA	72240
ATATTTGTAA	TAGCTGTAA	CTTTATCATT	GTCAAAATA	AAAGGATAAG	CTATTACAAA	72300
AGACAGCGGC	AATTGAGGTA	AAAGACTTCT	AATTCCAGTT	CCCCAGCTAA	AAGCAAACT	72360
ACTAAAAGGT	CTAAACAAAG	AATTTCTTG	CCCTTCTAA	GAATAGGAAG	CAAATCTAT	72420
AAAAAAGCA	TCCCAAAC	AAATATTTT	TAACAAAGGA	ATAGATATCT	GCACAGTATT	72480
TACAAAAGAA	CTGTAAATAT	TTTTCAAAAT	CCCCAACCT	CTAGCCTGCA	TAAATTTTC	72540
ACTAAGAATT	ATGTGGTGAT	GGGGTTGAAT	TTCAATTTCA	AAACCATTAC	CAAGAGGAGG	72600
TAATATATTT	GAATAGACAC	TCCTTAAGGT	CAAAATAATA	TCAAATAAG	GAGTAAACAC	72660
ATCCTCATAT	CCCAAAGAG	AAAAATATCT	CTCAAAGTT	GTAGAAGATT	TAATAAAATG	72720
GCTCTGACCA	AATAAAAATC	CACCAAAAA	ATCAAAGTT	TGCTTAAGTA	AAAATCCATT	72780
ATTAGATAAA	GAGGTAGAAT	TTCTTGATATC	CCAAGCCGCG	CTCAAAGTAA	GAGAATTTTC	72840
AAATCTAAAA	GTTTTATAAT	TGTCTCTTAA	ATAATAATTT	GAAGGTCTGT	TAACCTCATT	72900
ATCATAAAAA	ACATATTTTA	AAGCAGTTTG	CAAAGTGCCA	AGAAGGGTTT	GTTTGCCAAG	72960
ATAATTAGAA	AAAGTATACC	CGGTAAACGC	TCCAAAAC	AGTTTaAGCA	AAGAATAATT	73020
CATAGCATT	AAATCGGAAA	AGCTTTTAGC	ATCTCGATAT	TCTTCCCAAC	TTGTAAATGG	73080
ATCAGGAAC	TCCCTCTTGC	CAGAAAAAT	AGGCCCATTA	ATATCCTGAT	AAGCAGTATT	73140
AACGGAATGT	GAAAAATCTA	TAAATCCACC	TACGGTCCAT	CTTTTTTGAA	AAAACCAATT	73200
ATCTCTAAAT	GTCAAAGTAA	GACTTTGCTC	TAAAAAGAT	AAATTTAGTC	TTGCTGCAAA	73260
ATAATAGCCT	TCGCCTAAAA	AATTAGAAAG	CTCCCACTGC	CCAAATACTG	AGAATGGAAA	73320
TGAAGAATTT	GAATTGCCTC	CAAATTCAT	ACCAAATCCA	AAATTACTTG	TTGCTCGCTC	73380
CTCAATGTTT	AAATTTATTT	TCATAAGCCC	TTCTGTATTG	CCTGGAACAA	TATCAGGAAT	73440
TACATTTGAA	AAATAACCAA	GCTGCTGTAA	ATTTGCCATA	CCCATCTTAA	ACTTGTCCAA	73500

ACTAAAAACA	TCTCCCTCTT	GAAGAGGAAT	CTCTCTAAGT	ATTACATGCG	AAGCTGTATT	73560
TTTATTTT	GAAACAGTAA	TAGACTCAAT	ATGAGCTTTA	TCCTTTTCTA	AAATTTTAAT	73620
TAACAAATCA	ACAAATTCCC	CTCTTATCTT	TTGCGAAGGA	ATAATTTCTG	TAAAAATATA	73680
CCCTTCTCTA	AAATAACTTT	CCTTAATTTT	GACAAAATCC	TGCTCAAATT	TAGAATCATT	73740
AAAAATATCA	CCTTCGCTAA	AGGTAATAAA	ACTTTTTAAT	TCTTCCAAAC	TAAAAACTGA	73800
ATTACCAGAA	ATTTCAAGCT	TTCCAAATCT	AAAAACATTG	CCTTCTGAAA	GAAAAATATT	73860
CAAAAAAACT	TCCTTTTCTA	GTCTTTTAGA	ATCTTTAAGG	GAATCTTTAA	TATCAACAGT	73920
GCTATTGATA	ATCTTAACAT	CAATATATCC	ATTATTTT	TAAAAAGACT	CTAATTGACG	73980
CTTGTCTTTA	TCAACATTAC	TTTTTAAATA	TTTACCATCT	GAGAAAAGAG	ACACTACTCT	74040
TGATGCTAAA	GATTTCTCTA	AGGTACTGCT	TTTAAAGCTT	AAATTTCTCT	CAAAGTCAAT	74100
CCCCTTAACA	ACATATTTGG	GTCCAGCTAC	TATATTAAAA	ATAATATCAA	CTAAATTTCC	74160
TTCTTCTTTG	ATTTCAAAAT	TTGCAGAAAC	CTCAAGATAT	CCCATGTCTT	TATACATCTC	74220
TTCAAGCTTG	CCAATACCTT	TATTAACACT	TGCAAGATTT	AAAGGCTCAT	TGGTTTTAAT	74280
ATTCACCTTC	TCAACAAGTT	CGCTATTCCA	AAAAACTCTA	CTGCTATCAG	AAAAACAAC	74340
AGAATTAAC	AAAGATTTTT	CTTTTACAAT	AAATGTAATA	AAAAGATCCT	CACCATCTAT	74400
TTTAAATATA	GGCTTAATAA	GCCCAGAAAA	ATAATCAAGA	GAATAAAGAT	CAATTTGCAA	74460
TTTATCAAAA	ATTTCAATTAG	AATATGACAC	GCCAATGTAA	GGTTTTAAAA	TATTAATAAA	74520
ATCTCTCTCC	TTCTTATTCT	TAAGTCCTTC	AAAATTAATA	CCCTTTATTA	TTTTCCCCTT	74580
GTAATTTTCA	ACTTGACCAA	AACTAAAAAC	AACAAAAAAT	ATTAAAAAAC	TTACAAAAAA	74640
CAAACCTCTA	ATTGAACCCA	TCTTAACCTC	TTAACAATTT	AATATTTAAA	TTTCCAAGAA	74700
ATGCCTATAT	TATTTCTCTAT	TCCATCTAAA	CCTTTTTTCA	TAAAATTGTA	ATCAAACCTCA	74760
TAATTAACCA	AAAAAATGG	AGAATCAAAC	TCAATACCCa	AATTGACAAC	aAAATTTAAA	74820
TCTTTTGAAA	AAGGAGACAT	TTGTTCTTTC	AAAAAACCAA	AGCCCCCACT	AATAAAAACA	74880
CCCTCAACAA	GATATTTGCC	TACCTTAACA	CTTGTATTGT	CAAGAACATC	AACAAAAGTA	74940
GGATTCCCGA	TTTTGAAAAA	ATTACTATTA	ATAGAATTTT	TCAATATATC	TGTCTTTATA	75000
CTCAACAAGT	CTAAGTTTAA	TACAGAACGC	ATATAATCTT	CAATGGGTTG	AATTAAAAAA	75060
TCAAGAGCAA	TGTCACCTAC	TATTCCAATT	GCCATTTTCA	CAGCATTAGT	CCCTGCCGAT	75120
CGCAATCCTC	CCTCATACCC	TCCTATTGTT	GAGCCTGAAA	GCAAATATTT	AATTTCTTGC	75180
TCATTTCTAG	AAGGATAAGA	CATAAACTCA	ATTTTCCATA	AACTTAAAGG	ACTATCAATG	75240
CTTATTGTAA	CAAGCAGTTT	ATCATTCTTA	TCCTTAATAG	TATTTGTAGC	CTCCGCTTTT	75300

ATCCATGGAT	CAAATTTAGC	CCTACTCTCA	TTAAAGGATA	TATAAGAGCC	GCTTTTAAAA	75360
ATAAATTTTT	TATTATTATA	ATTAACAAAA	CCACTTGCAA	TATTCAAATC	TCCCTTAATA	75420
ATAAAATCAT	CTGTTTTTGT	GTCAGATTTT	ATTACAAGTT	TATCCCCTCT	TGAAATAGTA	75480
GCTTGTAATA	AAGAAATATT	ACTATCTGGC	CAAtGAAAAG	TAACACCGCT	GTCAAAATTT	75540
ATCTCAAGAT	CAGTTAAAAT	ATCAAAATCT	AGAAGATTAA	TATCAGTTTG	CAATCTTTTT	75600
GCTCGTTTGA	AAGGATTTAT	TAATAAATCA	ACAACAGAAC	TTTCAAGAGA	ATAAACCCAA	75660
GCATTTGAAA	TATTTAAAAT	GCCTTTAAAC	ATAATTTTCA	TATCAATTTCC	TTCAATTGAA	75720
AAATCGCCTA	AAGCATAGCC	TGTAAAGCTT	AAGGCAATTT	TTTCAAATTT	AATAGGAACT	75780
CCCGTTCTAC	CAGTCACATT	AATATCTATT	TTGTAATAAT	CAATAATAGT	ATCACTTAAA	75840
AAATTTAAAT	TTAAACTAGT	AGAAACAAAA	ATCTTGGAAT	ATCGATCTAG	GTAAACTCA	75900
TTACTAAAAA	TGATTTTATT	ATCCTGAATT	GCAACTGGCA	TATCAAATAT	TTCTAAAGCT	75960
CTATCACCAC	CAAATTTTCT	AGAAGCTCTT	AAATACTCAG	TGCTAATTGA	TCCTTTTTGA	76020
ATATTTAAAC	TTCCATTAAT	ATTAGGATTA	TACAAATCCC	CATCAATATC	GAATTCACCA	76080
TTTAAACAA	GATCGTAAAG	AATAAAATGA	TTGTCAACAT	TAAATAAAGA	ATGTGAATCC	76140
AAAAAATCTT	TGGTGATTAT	TTTTGAATCA	AATTTAATAT	CTCTCACATT	TCCAAGAATT	76200
TTATTTTTAA	TTATTTTGCC	TGACGAATTA	AAACTAAGGG	GCAAATAATC	TTTTAAAATA	76260
AAGCTATATT	CCCCACTGCC	ATAATGATAT	AAAACATTAA	CAAGGTCATA	ATCAATTGAA	76320
GCCATATTAA	ATTTTTCAAA	ATCATTGCTA	AATTCAATCG	TAAGATTGGA	TAAAGATTCA	76380
TTTTTATACT	TAATGTCTCT	AAAGAAAAAT	TCGCCCTCTG	TTTGGTTTT	CAAACTCCA	76440
AATTGCTCTT	CAACTCCACT	TTGGAAATTT	CTAAAATTTG	CATTAAACCT	GTAAGAAAAC	76500
AAGTCGCTAT	TAAAAACTAT	ATTAGATACC	CCTATAGAAC	TGCTTAGGTC	ATATCTTAAG	76560
CTTCCGGTAA	GAAACTCTTT	TTTATTGCGC	TTGGCTTTTA	TGTCATATAT	GTAAAGCTTA	76620
TTATCCAATA	ATCCTAAATT	CATAGAAAAC	TGAACGGGAA	CGCCCAAAAA	AGTTAATTTA	76680
TmTGCCTCAA	GATATCCACT	CAAAGAATAA	TTTTTAAAAT	CATTCTTTTT	AAAATTTAGC	76740
AAAAAATTAC	CATTAACCTT	TCCCTCTAAA	AGGGAGAAAG	AATTAAAATT	GTAAAAATCT	76800
AAGTCTTGAA	ATTTAAGCAA	AAAATAACTT	CCATATTCAT	CGGAATTAAC	CCCTAAAAAA	76860
TATCTCTCTG	AATTTAAGCT	AGAAAATAGA	TTCAAACCTC	CATTAAAATT	ATCTTTTAAA	76920
TTAAATTGTC	CATATCCTTG	CAAATTTGAA	AGTTTGTTTA	TAAGTTTAAC	ATCAGAAATG	76980
TACAATTTAT	CATATTCATA	TAAGCCCTTA	AAACCAAAGT	TGAAATTATA	GGCAGGTGTT	77040

TTAGAACTT	TTATTATATT	AACTCGTCT	ATTTTAAATT	TCAAATCTTC	ATTAAGGCCT	77100
AAATAGTTCC	CATTAAAATT	TATATTTATT	AAAAGATCAG	AGGTTTTATT	GTAAACCTTA	77160
AAATCATTAA	CATTAAGGGT	ATAAACTACT	TTAGAATCAA	AATAATTAAA	ATCTAATTTA	77220
ATACCAAGAG	GAGATTCGGC	AACAATAAAC	TTATCTTTAA	GATTAACATT	AAAATGCAAA	77280
GGATAAACTT	TATTCAAATA	AGAAAAATTC	GTGCTAATAT	TAAACCCACT	TTCAAATAGT	77340
TCAACAAACA	AGTTAGAATG	CAAATTGTGA	TCATTGTACT	GCAAATCAAA	ATTGTTTGAT	77400
TTGTAAAAAT	TTTTTCTCC	ACTGGCATCA	AGCACAAATT	TGAAATTGTC	TAACTTTGAA	77460
AATACCATGA	AATTGAATTT	TTTTAATTTA	TTTTTATGAT	AATTAAATTT	ATTAAAATTA	77520
AAATCAGAAA	CCAAATTTAA	ATATTTACCT	GAAAATAAAG	TCTTGGGAAA	AAAATTTATC	77580
AAGTGAGAAC	TGGGAATAAC	TTCTTTTAAA	AAAAGCAAAG	GAAATTCCTT	AATACCTAAG	77640
CTTAAAGAAA	ATTCTTCATC	ATTAAGATCT	CCTTTTAAAG	AAATTTGAGA	GTTTTTATTT	77700
TCTAAATAAA	TCAAATAGTC	AACAAAAATT	TTATCCTTTA	AAAAATAAGT	TTTTAAAATT	77760
AAATTTTGAA	AACTAAGATT	TCCAAGACTA	AAATTATCAG	ACTTAACCGA	AAAAATATCT	77820
TTATCTTTAT	TAAAATCTAA	ATACCCATTT	AGGTCGTAA	AGTTTAAAT	TTTTGCAGAC	77880
TTAAAGTTAA	GACGTCCCAT	TGGAAGCAAA	TCTTTCAAAG	AATAATAACC	TTTATAGTTT	77940
ACAACCCCCC	TTTTAAGCTT	TAAAAAAGCA	TTCTTAACAC	TTACAATTCT	ATCATCCCCC	78000
TTGATTTCAA	GCTGCAAGCC	TTGAACTTCT	TTTCCTATAG	TATCTACATT	TAAAGATGAA	78060
TCTATTATTC	CTGCATATCT	TAAATCTTTG	TCCTTAAAT	CATAAGAAAA	TGCCAATTGC	78120
CCATTTAAAC	TTATATCAAA	ATAATCTTTA	TAAATTTCAA	AGCCTTTGTT	TAGCTTAATC	78180
CAATCTAAAA	GACTAACATT	GAAAAATAAA	GCATCTAATC	GAACAAAACC	ATTAGCCTTG	78240
TCATAACTTA	AATTAAAATC	AAAATTCTCT	CTTCGTAAAT	TAAAAATTTT	TAAATTTCCCT	78300
TTTGAATAAT	TTATTTGGAA	CCCCTGCTCA	AGTAAAGAAA	AATAACTTGT	TTTAAATTCA	78360
AAAAAGCTAA	AATTAACATA	GCCATCTTCA	AAGCCTTTTT	TGAATTTCCC	CTCAAAATAG	78420
AAAGTTGAAT	CCAAAATTCC	ATCATCAACT	CTTTCAAAGG	GTAAATTAAT	TTCTAAATTT	78480
TTAACAGCAC	TAAAATCAAC	TACAGAGCTA	AATAAAAAAT	CTTCATCTAC	GGTACTTAAG	78540
GAAAAATTTT	TAACTTGAAA	ATTAAGCCAA	CTATTATCAT	TAAGCTTGAT	ATTAATATTG	78600
ATATTTTCTA	AATTAATATT	TAATCTATAA	AGGTAGTTTA	AAATTTTATT	AAAAACTGTA	78660
TTTTCATTGT	CAGAATAGGC	ATTGCTAGGA	TTTAAATCGC	CAGATAAACT	AAAGTCGTTT	78720
ATATCAAAAT	TGAAATTACT	TCCTTTAACA	TAAACATTTA	AAATAATATT	TTCATCACCC	78780
AAAATTAATT	TAAACAGATT	TAAATCTATC	CTAACAATAT	CTATTAATAT	TTTATCTTTT	78840

CCATCCAAGC	TTAACTCTAA	ACCGTCTATC	TTGATTGATG	ATAAGAAATA	CGGTGAAATT	78900
TTATCATATT	TAATCTTAAA	GCCAAATTTT	GATTCAAGAT	ATTTTATAGC	AAAAAACTTT	78960
GCAGAATAAA	TTTGAGCTTG	AACAAATAGA	TTAATGGAAA	AAATTATTAA	AACAAAAATA	79020
AAAAATGGCA	AAATCAACAA	TATAAATGTC	TTACTTCTCA	AAAACAACAA	ATTCATACAC	79080
TCTATCGATA	ATTATTATTA	TATAATAATT	ATCGATAACC	TAATTATTGA	CACCAAAAGA	79140
AAGGAAGAAA	AAATATTTGT	GATTAAAATA	TTGAAAAACT	TTTATTGCAT	AGAAGGAATT	79200
GATGGAAGCG	GGAAAAACAAG	CATCACTAAT	AAACTAAAAG	CTCTTTGCAA	CGATGAATCA	79260
AGGTATTATT	TTACAAAAGA	ACCATCAAGT	GGAATAATTG	GAGAAATGAT	AAGAAAGCAA	79320
TTAATGAATT	TTGAAAATCC	TTTAGAAGAA	TCAACATTTG	CATATCTTTA	TGCTGCAGAC	79380
CGACACGATC	ATTTATATAA	AAAAGGTGGA	ATACTGGAAA	TTTTAAACAC	AAAATCTAGA	79440
AAAATAATAA	CTGATCGCTA	TTTATTCTCA	TCGATTGCAT	ATCAAGGAAA	ATTAGGATAT	79500
GAATTAAATA	AAAATTTCCC	ATTGCCTGAA	AAAGTATTCT	TTATCGAAAC	AGACCCAAAC	79560
ATAGCTTATG	AAAGAATACA	GAAAAATAGA	ACACAAAGTG	ATCTTTTTGA	ACTTGAAAAA	79620
TATAAACTT	TTGAACAAAT	TGCTCTAAAA	TATTTAAAAA	TATTTAAAAA	ACTAGAAAAA	79680
AAAATTAATG	TGATTTACAT	CAACAATTCA	ATAAAAAGATA	ATTTAGATAA	AAACGCAAAA	79740
AAAATTTTCA	ATCTAATAAA	ATTCTAATAT	AATTAATCAT	ATGCATATTT	TCAAAAATGT	79800
CCCCTTCCAA	ATAAATTTAA	TTTTATTTCT	TTTAGTATCA	GTTGCAAAGA	TAAATGCATC	79860
GTCCAAATTT	TATTACGCAG	AACAATGGTA	TGTAATTTTT	AATTCTCAAA	TGAAAAAAA	79920
ACCTGAAAC	TATAAAAAA	ATATATTTTT	TCTTCAAAAA	GCCTTAAAAT	ACCCATTTGG	79980
AAATCCAAAA	TATTCTCTAA	CTAAAATAGA	AACCAAAGAA	CAGTGGGAAA	AATATAAACT	80040
TCTTTTCAAA	ATGCATGTAA	ACTTGCTTCT	AGTTAGGCAA	AATTTACATT	TAGGAGATTT	80100
ATTCGACACA	AGAAATTTAT	ATTTTTTCAA	AACTCCAGAA	AAAGATGGAA	TTATTTCCAA	80160
TCTAGAAAAA	TCAAAAAAAT	TATATAAACT	AGCTATTAAT	TACTACAGCG	AAGCACTAAA	80220
ATACCACAAA	AACTTGAAA	ATTACACAAC	TGTTAAACTA	GAAAACGATG	GAATAACAAA	80280
CTGGGAAGAT	GAATATCATA	AAATTTCTCT	TAAAGAGCTT	AATTACTATG	ACATTATTAA	80340
AAAAGAACTA	CTAAGAATTG	ACGAAACTAA	AGCATTTTTT	GAACAAGGGC	CAAACATTAT	80400
TTAAAAAAAC	TCTTTGCCCT	CTTTGGAAAA	AAAAATTTTA	TTTATATAAT	CCTTATTTAA	80460
AGAAAACTTA	AAAACAAGAT	CTTTAAAATT	ATCCTTACTC	AAAATACTAT	ATTCTGAGAA	80520
AAGAGTTATT	AAGGCTCTTT	CTGCTAAAAA	AGGCAATTCT	AAAATATTTT	TTAAAAATTT	80580

GGCTCTAAAT	TCACGCGCTT	TTTCACTATC	TAAAGAATTT	ACAAAATTAG	ATTTTCAAC	80640
ATCAGTGCTT	AGTGCTACAA	TTTCATTTTC	AAGAAATTTT	TTATCTCCTT	GTGAATATAG	80700
CAAAGCAAAT	GGCTTTTAAA	AAGAAAAAAA	ATATTCCTTA	CCTATAAGAT	AATGGTGCAT	80760
TGAACATCTT	GTAAGTAAT	TTGGGTAAAT	AGCAGATTCA	ATCTCTTCAT	CGCCACCAAC	80820
AATAAGCAAA	GGATCAAAAT	AAGGAGCTGG	AATTTCTTTA	TTGTTGTAAA	AATAATACCT	80880
ATACGAAAAA	AAAGATTCTT	CCATGCAATC	AATGTGCCCA	CAATAAGCTC	CAAGCTTATA	80940
AATCTTATCT	GAATATTCTA	CTAAAAGTTT	GGCAGTATCA	TTGAAAGATT	CTTTTCTAAA	81000
ATTATAAAGC	AGTGTAGGTA	ATATAAATAA	AATATTCTCA	TTTGATGAAA	TTTTATTTAA	81060
AACAAAAATT	AAACGTTTCAT	CATAAAAACA	TGCCTCATCA	ATAATAAAAG	TGCCACAAC	81120
AGGATTAGAG	GCTATTAAAT	TTTCAATATC	AAAAGAGTTG	CTAGCATAAC	CAATCTCATC	81180
AATTTTATCT	TTTCCACCGC	CTCTATATGG	TATTACGTTT	TCTGGATAAT	CTTGAAACCT	81240
CCTCTTGTCG	AGAAAATTTT	TAATAAAAAA	TACATTAACC	CTACTTCTAT	TTCTTTAAT	81300
AATATTGCCC	AATACCTTGA	AAGATTTTTT	TCTTACAACA	AGCGAATCTT	TATAAATTTT	81360
TGCAGCATAT	TCTGTTTTTC	CACTTCCCAT	GGGTCCAAC	ACAAGAATTA	AGTTTATTTT	81420
TACCCATAAA	TCAAAATGAC	TAACAGAGAC	AATGTTATTT	AATTTAGTAT	CTTCTTTATT	81480
AGCAAAGTCT	AAACAAAAAC	CCAAAAATCC	TCCCTAAAGT	AAATCAAATT	CAATTATATA	81540
AATAAAAAACA	ACAAAAACA	TTAACATTAA	AAGCCTAAAA	ATTAATAATT	TAGGATCTTA	81600
TTAAAGCTAT	TATTCAAAAG	AATAATAGCT	TTCAAAACTA	TCATCATCTA	ACAAAGCTTT	81660
CTTTATTTTT	AGTTTATTCT	TCTCATAAAT	TTCAATATAA	TTAAATTTTT	TAGAACTATT	81720
AACTTTTTTA	TAAATTGAAA	CCAACTTCC	AACCACATAA	TTTTTAATTT	CTACCAAATT	81780
AGATCCAAAA	TATTTTTTTT	TACTTATCAA	GCTTGATCTA	ATATCCAAAA	CAGAAGTTTT	81840
TTCAAAAAAT	AAATTTAAAG	CATTTGGAAT	AAAAAAATCC	CTAATATCAA	TCAAGAACAA	81900
CTCTCTGATC	CCAAAACAT	TAAAATTGCA	GCCCAAATTG	TTAAAAAGAT	TAAACTTAAA	81960
AGCATGTAAA	GCAAAGGTAT	AAACATCTCT	TTTATCAGGA	TATTGAACCT	CAATCATATC	82020
AACATAAGGA	TAAACAGAAT	AATTTATTTT	ATAACCAAGC	AAACTATTCT	CATAATAAAC	82080
TGGAGACTTA	TCTTTAAAAT	AAATTTTATT	AAAATATTTG	CTATTTAAAC	TAAATTTGCT	82140
GTCAAAATCT	ACTATTCCTG	AAAAAATGCC	TACTAAAACT	CCATCTTTTA	GTATGGCAAC	82200
ATTATTTTTA	TTATCAATGC	AATAAATTC	TGACAAACTT	TTATAATCTT	CTAAAAATTT	82260
ACTCTTCATG	TCGGGATCCT	CTAAAAGATC	ATAAACATT	TTATATTCAT	TTATTGTCTC	82320
AGGAGGAAAT	TTTTTAAAAA	TCCTATAAGC	TTCATCAGGA	TCTAAAAGTT	TGTATTTTAA	82380

AAAATAAATA	CCGTGAATTC	CATTATATTC	TTTAATTTTG	CCCTTAAGAA	AAGAAAACAA	82440
AGAAACAGAA	TTGTCATAAT	TAAAATTTGA	ATTTCTAAAT	ATATAATTAA	TGTTTTCCGT	82500
TGTAAAGAAA	AAATCTTCTT	TTGATCTAAT	AGTATTAAAG	ATGCTAAAAA	AAAACCTATC	82560
TTCTTTTAAA	AACCATTTAA	AAAATAAAAA	ATCATTACCA	TATAAGAAA	AAGCCTTAAA	82620
TAAAATTTTT	TTGAAATCAT	CTCCAAAAC	TTCAAGTCTT	CCAGAAGCTT	CAAGCCTCAT	82680
GCAAATTAAA	TCTTTGTCTA	TTTGCAATTC	TTCATTCAAG	TTTCTATACA	AAGTAATAAT	82740
TTCTTCATAC	ATATTCTTAT	TTTCATTTTG	TCTTAAAAGC	AACGAAAAAT	AATAAGTGTA	82800
TACTTCTTTA	AGGGTTACTG	CTTTAAAGTT	ATTAATATTA	ATTGCTTCTT	TTAATAAATA	82860
AATTTTATCC	GTAAAGCAA	CATTCTCTTT	TAATGAAAGC	TCATAAGAG	CATCTGAAGC	82920
TAGCAAATTG	AAATCAAGAG	ATCCATAAAT	TAAATCATTT	CCTTTTTGTA	AATTGGAAAA	82980
ATTTTGTAGC	TCATTGAAAA	GAATACTTGA	AATATCTTCA	TCTTCTTTGA	CTGAAATTGC	83040
ATATAGGTTT	ATAGAAATAA	AAAGAATGAA	ACCCATCTTC	AAATAAAACA	TAAACACTTG	83100
CCTAACCTTT	TACATCCTCC	CCCTTTACTT	CTTTTGTAGT	AGAGTCTGAT	GAAAATAAAT	83160
CGTATTCATC	CTTATTAGCT	TCAAAACCCA	AAAGCTCTCT	CACTTCTTTG	TCTGTCAAAG	83220
TTTCTTTTAA	AACAAGTTCT	TTTGCAAGCT	TAACAAGTTG	ATCCTTATGC	TTTAAAAGAA	83280
TATCCGATGC	CTCTTTTAAA	CATTCTTCAA	GTATTCTTTT	TACCTCTCTG	TCAACTTTAT	83340
CGGCAGTGTT	CTCAGAATAA	GCTTTAGCCT	TTGAAAAC	TTTCGGAAGG	AAAATAGGTG	83400
CTTCATCGTC	TACTAAAAAT	ATTGGACCAA	CTTCTTCACC	CATTCCCCAC	TCCGTAACCA	83460
TTTTTTTAGC	CAAAC	CTTGCATTA	AATCATTTTG	AACACCCGCT	GTTGTAACAC	83520
CCAAATTTAT	TTGCTCGCTA	GCATAACCAC	CATAGCATAT	CTTTATTTTG	TCAAGAATTT	83580
GGTGTTTGTT	TATTGAAAGT	CTATCTTCCC	TTGGAAGAGA	AAATGCAACA	CCAAGTGCCC	83640
TGCCCCCTGG	AATAATGGTA	ACTTTGTGAA	GTGGATCAGC	ATGTTCAAGA	TAATAGTGAA	83700
GCAAAGCATG	GCCTGCCTCA	TGATAAGCCG	TCTCAAGCTT	TTGCCTATCA	GTAATAGTCA	83760
TGGATTTTTT	TGCAACTCCC	ATCAATATTT	TATCTCTGGC	TTCTTCCATA	TCCTTCATTA	83820
AAATTTTCATC	TTGATTATTC	CTTGCAGCTA	TTAAGGCTCC	TTCATTAATT	AAATTTGCAA	83880
GATCAGCACC	ACTAGCTCCA	GGAGTAGCTC	TTGCTATTAC	TTGTAAATTA	ATATCTTTTG	83940
AAAGCTTCGT	TTTTAAAGAA	TGAATATTTA	ATATTGCCTC	TCTTTCCTTA	ATATCAGGCA	84000
AAGAACTGT	TACTTGCC	TCAAATCGTC	CAGGCCTAAG	CAAAGCAGAG	TCAAGAACAT	84060
CGGGACGATT	TGTAGCGGCC	ATAACAATTA	CATTGGTATG	CGTTCCAAAT	CCATCCATTT	84120

CAACTAACAG	CTGATTAAGA	TTTGCTCTC	TTTCATCATG	ACCACCGCCA	AGCCCCGCAC	84180
CACGACTTCG	ACCAACAGCA	TCAAGCTCAT	CAATAAAAAAT	AATACATGGA	GAATTTTTTTC	84240
TAGCATTATC	AAATAAATCT	CTAACACGAC	TTGCTCCAAC	CCCAACAAAC	ATTTCAACAA	84300
AATCTGAGCC	TGACATGTGA	AAGAAACTAA	CCCCAGCCTC	ACCGGCAACG	GCTTTGGCAA	84360
GCAAAGTCTT	GCCAGTACCC	GGAGAGCCCA	CTAAAAGCAC	TCCTTTGGGG	ATTTTTCAC	84420
CTATTTTTTTC	AAATTTTTTTT	GGATTTTTTAA	GAAATTGGAC	AACTTCTCGA	AGCTCTTGCT	84480
TAACCTCTTC	TTGACCAGCC	ACATCTTTAA	AGGTGATTTT	ATTCTTTCCA	GCTTCATACT	84540
TTTGAGCATT	ACTTTTCCCA	AATGTAAAAA	CCTTCCCACC	GCCACCTTGA	GTTTGACGAA	84600
ATATAAGAA	AAAGAAAATA	AAAAACAAAA	TCCATGGCAA	AGTTTGTAAT	AAAACCCCAA	84660
TCAGAGAAGC	TTGACTTTTC	CCTGAGCTAA	GCTCAACTTT	TTTATTTTTT	AGTTCTGAAA	84720
GTAAATTTAT	ATCAAGATAG	GGAATGCTGG	TAGAAAAATA	AGACTTTGCA	AAGTTAGAAC	84780
CCTTGACGAC	AAATTGAATC	AAATTTTTTAT	CAATTATTAC	TACAGACTCA	ACTAGACCAT	84840
TGTCTAAATA	ACTCTGAAAA	GTGCTATAAG	GAACATTTTT	ATAGCTTTCC	CCCCCCTTA	84900
TAAATATGA	CATAAATATT	GCTGAAATTA	GAAAAACAAC	AACAAGTCCT	AAAATCCAAT	84960
TTTTATTCTT	TTTTTTGTGG	TTAGATTTTC	CATTGTTATT	CATATTATTA	TTGCCATTCA	85020
TTCTTTTAA	AGCCCTCCAA	TCAAAGATAT	ATTAATTTTT	TTAAGAATGC	TTTTTTCACT	85080
CCATACTAAG	TTTAAAGTAT	TTAAATCAAT	AATCCCAATT	AACCTGTTAT	CTAATGCTAA	85140
CAACATTAAA	TAAGCCGGAT	TACACCTTAT	AAACTTAGAA	AAAAATTTTT	TTGCTTTCAA	85200
TCTATCCTTA	AAAAATTTAT	ACCTAAACTC	ATAAGAACAA	CATTTTAATC	TTGATACAGA	85260
AGCTGCATTA	CACTCTAAAT	ATTTTAGCAA	AATCTTACCT	AAAGATAAAC	TATGCCATTT	85320
ACCAACTTCC	AAAATAAAAT	CAAAAGGTTT	GTAAAATTTT	TCATCCCTTT	TAAAAATTAA	85380
ATTAATTTTA	TTATGCCTTT	TTTCTAAAAA	AAAATCATTG	GTTTTTAACA	AAACATTATT	85440
TTTTTTCCTA	TTAATCTCTA	CTTTAAACGC	TTCATTAAGA	GCTTTATAAG	AAACTTTGGC	85500
TGCAATTCCT	TCTGAATTTA	AAATTTTAAA	AATCAATCTA	AATACCAAAT	ACTTAGGAAA	85560
ATCTAAGAAA	GTTTTTCAGAT	CAAAAGAATA	ATAATATTTA	CCTTTCTCAA	CAGGAAAAAA	85620
TCATCTTTT	CCAAAATAAT	CCGCAAATTC	CTTTGAAAAT	TCAGATATTC	TTTTAAGACA	85680
TTTTTCATAT	CCTTTAAAAA	CCTTTTTTTAT	AGCGGGTAGC	AAATTATTTT	TAACCCTATT	85740
TCTTAGATAT	AAATTTTGAG	CATTTGTACT	ATCAACAAAA	AACCCAATAT	TATTCAAAGA	85800
TAAAAAATTT	TCAATTTCTA	GTCTTGAAAC	CTCAAGCAAG	GGCCTTATAA	TGTTTCTATT	85860
GACACTAGGA	ATACCTGAAA	GACCATCCAA	AAAAGATCCT	TGAAAAAATC	TCATAATTAT	85920

TGTTTCAATT	TGATCGTTTT	CGTTATGAGC	AAGTGCAATA	TAATTTGCGC	CATTTTCTTT	85980
AAGAGCATTT	TCTAAAGCAA	TATATCTAAA	TTTCCTTGCA	AGCTCCTCAA	TAGATACTCC	86040
CAACCTAGCT	GATTCACCTT	TTATATCTAT	ATCACATTTT	TTAATTTGTA	AAGCAATATT	86100
GTAAAGATCG	CAAAACCCCT	TTACATGCTC	TATCTCTTGA	TTTGTCTCAT	TATCAGATCT	86160
AATAAAATGA	GCAAAATAAA	ATGCAATAAC	ATTATTGCTC	AGGTAATATT	TTAAATTCAA	86220
TAATAAAGCC	GTAGAATCAG	CCCCGCCAGA	AAAAGCAACA	ATAACCCGAT	TTTTATCTAA	86280
AGAATTTTTT	TTATAAAATT	TATCTATTTT	AATTTGTATA	TTTTCGTCTA	AAAAATGCAT	86340
TAAAAATTAC	TAGAATTAAT	TTTTTGAACC	TTATTTTAA	AATTTGCTTC	GCTAATATCA	86400
ATCAGCTCAT	CACCTAAAAA	ATCATACAAT	TTTCAAGTC	TATCCATTTT	ATCCTTGCAA	86460
AATCTAGAAA	GAACAAACT	CTTTATATCA	CGCATAAGAT	TGCTTCCAAC	CCCAATATAA	86520
AGTCTACTAT	AGTTAGAGCT	TCCAAGAACA	CTTGAAATCG	ATCTAAGACC	ATTATGAGTA	86580
GACACACCAC	CTCGCTCTTT	AAGCCTGCAT	TTTCCCAGGG	GAAGATCAAC	ATTGTCAAGA	86640
ACCACTAATA	AATTCTTTAT	ACACATATAA	AAATCTGAAA	ATATTGAAGG	AAATAGAGAG	86700
CCACTTAGAT	TCATATAGGT	CAATGGCTTA	ACCAAAATTA	CTCTTCCGGA	AATCATTTTT	86760
AATTCTGAAT	ATTCATACTT	TTTTTTTCTC	TTTAAAAAAA	GGCCGTTTTT	AGAAACAATT	86820
TTATCCAAAA	GAGAAAAACC	AACATTATGC	CTAGTTAAAG	AAAATTCTAA	TCCAGGGTTT	86880
CCTAATCCAA	GTATCAACAA	CCCCATAAAA	TCACTTTATA	ATAACAACCT	CAAATTCTC	86940
GTCATTTTCT	GCAAGCCTAA	CATTAGAAGG	CAACACAAGA	TCTTTAAGAA	GAACGCTATC	87000
TCCTTTATTA	ACAGGAGTTA	AATCAAGCTC	TATGAATTCT	GGCAAATCCA	AGGGCAAAGA	87060
CCTTACCTTA	ACCTGCTCTT	TTAAAACAGT	CAAATTTCCA	CCTTCTTTAA	CCCCAATAGA	87120
AGCTCCAATA	AGCTTAATAG	GAACATATTT	TTCAAGCTCA	ACATTCCTAT	CTACTTCGTA	87180
AAAATCAATA	TGATAAATGA	GCTTGCTGGC	AATATTTTCT	GCAGCATCTT	TAACAAAAAC	87240
ACATCTTTCA	AGCTTGCCAT	CATCTAATAT	TAAGACAGTA	TTATCCGTAA	ATTTTGCAAA	87300
TTTCTTATTA	AATTCACCTAC	TCTTAATTCT	CAAGTGTGAA	ACATCCTTAC	CTTGCCCGTA	87360
AACAACAGCT	GGTATTTTCA	ACTTAGCTCT	TATTCTGCGA	GCATTAGAAG	ATCCAAAGCT	87420
TGATCTATAT	TGGCAACTTA	AAACCCTGCT	ATTTTCCACC	AAAAAACTCC	TAATGTAAAA	87480
TACTGGGACG	GAAGGATTCG	AACCCTCGAA	TGACTGGACC	AAAACCAGTT	GACTTACCAC	87540
TTGTCGACGT	CCCAAATACC	TTATTAATAC	TTACAATTGA	CAAGCATACA	AGATACAACG	87600
AGCTTTGTCA	ACAATTATCT	AAAAATCTAA	TTCAAGTTCA	AGATCGGCTG	GATTTTCTCT	87660

TATATATTCC	TTAAAAATAG	AGTTTGCAA	AGCTTTTCTA	AAATCCTGAC	TAATAGGATG	87720
TACAATGTCT	TTATATTAC	CGACTCTAGT	TCTTCTGTTA	GGCATCGCAA	TAAATACTCC	87780
CTTTTGCCCT	TTAATAACTC	TAATATTGTG	AAGAACCAAA	CAGTTATCAA	AAGTAACTGC	87840
AACATATGCT	AATAATTTAG	AACCAGAATT	TTTACTATCA	ACTTTCTTAA	TCCTTATGTC	87900
TGTAATATCC	ACTTATAAGC	CTCCCGCAAA	AAGTACATAA	CTTAAATCTA	AAATATTTTC	87960
TATTTTTTGT	AAACACGTTT	TTATGATATT	TTTAGTTTTT	TTAATTAATT	TTAATTAAAC	88020
TAAGTAATTA	GGGATAAATA	ATGGTTCCTT	TGGACCATAT	TTTTTCAAGC	TCATAGTATT	88080
CCCTAGATTT	TTCACTCATT	AAATGAATTA	CCAAATTTGC	ACCAGAAACA	ACAGTCCAGT	88140
CATAAACCAA	CCCTTTTCCT	TCAGCATTA	GATTAATTTT	TTTTTCTTTA	AAGAATTTAA	88200
TTATCTTGTC	AATATATAAA	GCTTCCATTT	GCTTAAATGA	TACAAAAGTG	GCTATTATAA	88260
AAAAATCAGT	CCAATTACAA	ATATCGCTAA	CATTAATGCC	TATAACATCA	ATTCCATTAA	88320
AATCACTTAT	TATTTTACAT	AAATCATTAA	TATCATTTAC	TTTAAACATA	CCTTCCATCA	88380
AAATCATCTC	CTAAAATAAC	TATTACATCA	GGATTAATAT	CAAGATTATC	AAGCTCTAAC	88440
AATCTTTTAG	CTTGAACCTC	AGAAATTGGC	TTAATATTTG	AAGTTTAAAT	TACCTCTCCA	88500
ACCCTAACAG	CCATTTCTAA	ATTATCCGAA	TTATTTATAA	TCAGGGTATT	TTTATAAGAA	88560
TTTTTATCTG	CATTACCAA	TTTTAAAACT	TTAAATTTTA	AAGAATTAAA	AATATTTGCT	88620
GTTTTTTTTG	CAAGCCCAAC	AACTTTTGTT	CCATTTAAAA	CAACAATCTT	TACTATCTCT	88680
TCTGCACCTT	CATTAACCAA	CTCTTTGTTT	AATTTATCCA	CCGATTCTTT	TAAAATAGCA	88740
CCCCCATAAT	AAGGAAAAAC	CACCTTTATC	AAATTATTAT	CATTATCCTT	AAAAATCTCT	88800
TCTTGTCCCT	TAATATTAAT	AGAAATAATT	TTATCATTAT	TTATTTTATA	ATTTTAAACA	88860
ATATACTTAA	AAACAACCTC	TGAAAGGTTA	GTATCTAACA	TGGAATATAT	TTTAAAAAAA	88920
CTGTCATTTT	CAATGCCAAA	ATCTGAAATT	TGAAAAAGAA	GTCTTTTAAA	AAATTCCTTA	88980
AAAAATTCAA	CTCTCTCTTC	AAACTGATTA	ACATCATTA	AATATCTCAA	ATAATCATAA	89040
GCCTTATCAC	CATCAAATTT	AGAAGTGCCA	GAGGGTATTA	AAATAGAATC	CTCGAAACTA	89100
TAAACTTTCA	CTGGGTTTTT	AACAAGAAGT	CTAACTCCCC	CTAAGTAATC	AATAAGCCTA	89160
ACAAAATTTT	CTTTTTGAAA	ACGAATATAA	TAATCTGATT	CATGAGATAA	TTGTGTATAA	89220
ATTTTAGATA	AAAATTTATT	AAAAGAATTT	TTTTTATAAA	GATCTTTAAA	CCAAGATATA	89280
TTCCCTTTTA	AATCTTCATA	TCCAGTATGA	ATTGGAATAT	CAAAAAACCC	AATATTTCCCT	89340
GTTTTTCATAT	TAATAAAAAT	TTCTTGCATA	CTTACAAGGT	TTTTGTAAAG	ATCTTCTATT	89400
AGAAACAAAA	AACTAATATT	ACTCTTTGTA	TTAAGCTCGA	AGTAAACCAA	CTCTTTTTTC	89460

GAACCTCTAA	TAAAAAAAT	TACTACACTT	GCTATTATTA	AAACAATTAA	AAATAAAAAA	89520
ATTAAATCCT	TTCTCAAACA	TTTACCTTTT	TAACATATAA	ATTATTATCT	TTAATGTATT	89580
TTAACACACC	AAAAGGCAAT	AAATAGCTGA	CGGGCAATCC	ATTTACAATT	CTATTTCTAA	89640
TCTCTGATGA	GGAAATCGGT	ATTATTTTAT	TATCTATATA	AATATGCTTA	AAAGAACTTT	89700
TAAGTCTCTC	TTTGTAGATT	CTATGAGCAA	CAACAAGTTC	AACAGAACTT	ACAATACTTT	89760
GAGGATCTTT	CCATGAATCA	AAATTTTGAA	AAAGATCATC	GCCAATAATT	AAAAAAAGTT	89820
TATCGTTTTT	GTATTTTTTT	TTAACACAAG	AAATAGTATC	AACAGTATAA	GTTATACCAC	89880
CATTTATTAT	GTCGCAATCA	TCTATGAACA	TTTTATCTTC	ATTCTCTAAT	GCAAGCTTGA	89940
GCATATCTAT	TCTATTGCTA	ACACTAACAT	TCTCATCAAT	CAATTTATGA	GCTGGATTGC	90000
AAGTAGGAAT	AAATATTACT	CTATCAATAT	TTAATAAATA	CTCTATTTCT	TTAGCCAAAA	90060
AAATATGTCC	AATATGAACT	GGATTATAAG	TGCCCCCTAA	TATTGCAATT	CTCACGATTT	90120
CTTTCCTAAA	TAAAATCTGA	TATCCAAAAA	CCAAGTCTTA	AAAATAAAAA	GCCCACAATA	90180
AAAATCAATT	TTATTAAAAA	GTTTTAGCCA	AAATAAAAAA	TTCTTTAATA	AGTTCATCAA	90240
TTCTCTATT	CTCATAAATA	GAGATGCCAA	CAACCTTTTC	TTTTCTAAG	GCTTTTATCA	90300
GGCAATCAAA	ATTTTTCTCA	GAACCGTCCA	AATCAAGCTT	GTTGGCAATA	ATAATTTTTT	90360
TTTTATTAAA	AAGCTTATGG	CTATAAGATT	TTAATTCATT	TAAAAGAATG	TTATATGACT	90420
CCAAAAAATT	TGCTTCAGAA	ATATCAATAA	CCAAAGCTAA	AATTTTAGTT	TTAGCAATAT	90480
GCTTTAAAAA	TTTAGTCCCG	AGCCCTACTC	CAAAACTAGC	ACCTTTAATT	ATTCCGGGAA	90540
TATCTGCAAT	AATCAAATCA	TCATAAGAAC	GCCTGAGCAT	ACCAAGATGA	GGAATCTTTG	90600
TTGTAAAAGG	ATAATTTGCG	ACCCTAGATT	TTGCTGAGGT	TATCCTATTA	AGAAGAGAAG	90660
ATTTACCAGC	ATTGGGTAAT	CCAACAAGCC	CAATATCCGC	CACCAAAAAA	AGTTCAAGAC	90720
GCACGCTCAA	ACTATTACCC	GATTCTCCAG	GTTGAGCAAA	CCTTGGAACC	CTTCTAACTG	90780
AAGTTTTTAA	ATTCCAATTA	CCAAGACCCC	CTCTGCCACC	TTTTAAAACA	ACAAATTTCGT	90840
CATTTAAATT	TTTAAGCCTA	TACAAAAGAG	TTCCATCATT	TTCAATTATA	ACTTCTGTAT	90900
TTGGAGGAAC	AAAAAGAGTT	AAATCTTTAC	CATTAGCACC	ACTTCTTTTA	AAACCCATTC	90960
CAGGTTTACC	ATTTTCAGCA	CAAAGCACAT	GACCATTTTT	GTAAAAAGAT	AAAGTGCTAA	91020
GATTTTCCCT	CACCTTGAAA	ATTACACTCC	CACCACTCCC	ACCGTTTCCG	CCATCTGGAC	91080
CACCTTTTGC	ATTAAACTTT	TCTCTTAAAA	AAGAAACACA	CCCAGAACCA	CCATTGCCCG	91140
AAACTACCGT	TATATTTACA	GAGTCCTTAA	AGTTATACAA	ACTTTCTCCA	ATTTTTCAAT	91200

TAAAACCAAA	AATCTCCAAT	CTTTTCAATT	AAAACATAAC	AATACTTACG	TATTTTCGCC	91260
CCTTTTAAAGT	TTTAAACTCT	ACCTTACCAG	ATGAAAGCGC	AAATATTGTA	TAATCTCTTC	91320
CAAGACCAAC	GTTTTTACCT	TTATGAAACT	TTGTACCTCT	TTGTCTAACA	ATTATCTCTC	91380
CAGCTTTAAC	AAACTGACCA	CCACTTCTTT	TAACTCCAAG	TCGCTTGGAT	ATAGAATCTC	91440
GTCCATTTTT	TGAACTACCA	CCACTTTTAC	TTGTTGCCAT	TAATTTTCCT	CCAAAATAA	91500
TTTAATATCA	TTAGGATACT	CAAAGACAA	ATCATTTATG	CCTCTAATTA	AAAACCTACT	91560
ATAGTAAAAA	AGACTTTCTT	TGTTCAAATC	CTTAAAAAAG	GGCTTAAATT	CTAAATAACC	91620
TCTTTTTGAA	TTTTTCACAA	CAAAGCCTC	ACCCTCAAGA	TCAAGAACAC	TAAAAAAGGT	91680
TCTCAAAATA	AAAGAAAAAG	AAGAACAGAC	AACGTTAACA	TTATCTTAC	CTATAGCATG	91740
ACCATTGGCT	AAAAGATAAA	TAATTACATC	GTCTTTTACT	TTTACCACAA	CATTAATCAA	91800
TATACTTAAA	AAACTATTTT	ATCAACCAAA	ATATAAGAAT	AGGTTTGCCT	GTGCCCCAAT	91860
TTTCTCTCAC	TTGATTTTCT	TCTTCTGTAT	CTGTAAGAAA	CAACCTTTTT	ATCTTTTTTTA	91920
TCTTCTTTAT	AGGTACATCT	AATAAGAGAA	TTTACGACAT	AAGGCTTTCC	TATTTTAACC	91980
TCTCCGTCTT	TATTAATAAG	CAAAACACTA	TTAAATTCCA	ACTTATCTTT	TTCAACAGGA	92040
GAAATTTTGT	CTATTTTTTA	AAATTCACCC	TCAATAGCCT	TATATTGCTT	GCCATTTATT	92100
TCTACCAGTG	CATACATATA	TTACCTCAAC	TAAACATTGT	AAATTTAATA	AAAAAGAAAA	92160
GTCAAGCATT	AAATTAATTA	TACCTTACAA	GGGAATTGAC	TTTATAACTT	TCAAGACTTT	92220
GTCTACCATT	AATAGCGCAA	AGCTCAATAA	AACAAAAAAT	ATCCTTAACC	TTGCCCCCGG	92280
CTCTCTCTAG	CAAATTTGCA	GACGACTTTA	AAGTTCCACC	GGTAGCTAAT	ATGTCATCTA	92340
TTAAAAGAAT	ATTGGAATAC	GTCCTAACAT	CGTCTTTGTG	CACCTCTATT	CTCCCAAAAC	92400
CATATTCAAG	CTCATACTCT	TCACTAAAAA	CCTCTCTGGG	CAATTTACCC	TCTTTTCGAA	92460
TTAAAACAAG	GGGTAGCTGC	ATTTTAAAG	ACAAAGGAGC	ACCTATTAAA	TATCCCCTAG	92520
ACTCAACAAC	TGCAATGCAA	TCGATCTTTT	TAAAATTATA	AAAAGAATAT	ACTTCATTTA	92580
TTAATGAACT	ATAAACTTCG	GGTTTTAGCA	AAACGCTAGT	AATATCATAA	AAAAGAACAC	92640
CCTTTTTTAGG	AAAATTGGGT	ATTTTGAAG	TAAACTGATC	ATAATACTCT	GTCTTATTTT	92700
TCATAACCTA	CTTATCCCAT	CTATATGTTT	ATATATTTTT	CAAATATCAA	GACCAAAAGC	92760
CAAATTTATT	TTTTTCACAGG	GGGCGCTGGA	ACAAGATTAT	GCAAATCTAA	TTTTGGTATT	92820
TCATCATCTT	TAGCTCTTGT	CCAAATTTTA	CTTCTACCAA	AAATCCAAAC	TTTCCCCTTG	92880
GTAATAAGAT	TTCCGGTTTT	ACTATCAACA	CGCATCTCAG	AATTATAAAT	TTTACCGTTT	92940
TTAGGATCTA	TTATTTTGCC	CCTATCCAC	TTTTTAGAAG	AAGAAGAATA	CTTAAGACCC	93000

CACATAAAAT	CAAGACCCCTC	TATTGCAAGA	TTTTCAAACC	CAACTACAGT	ATCTCCTGAA	93060
GGATTTTITAG	CATCATACTT	TTTGCCATCT	TTTATTATAG	TTAAAATTCTG	GCCATAAACT	93120
TCCCCATTAT	ATTTATAAAT	ATAGATAATA	GAATTCTTTA	TGTTACTTAC	ATCATTATAA	93180
CCAACCCAAT	ATCCTAAAAC	TTCATTTTCA	AAAACAGGGT	TTTCATCCTT	GCTAACAATG	93240
TCCTTTTCAT	TTGAATCTTC	TGAATTTGCA	AATAAAAGCA	TTGAAAAACA	AAAAAAAAGA	93300
AAAAACTTTG	AAAAAATCT	AGTCATCAAT	CCTCCTTAAA	ACCAAATTAT	AGCTCTTTTT	93360
TTAAATTACT	CATGTAAGGC	AGCTAATTTA	AAAATTAGCC	ACAAACATCA	TTATACAAC	93420
TATTTTATAA	TTTAAATTAT	TAAGATAAAA	ACCTAGACAA	AAAAATATAA	AATTAAGCAA	93480
AAACACAACA	GGCCAATTAA	TTGTTATATG	GGACAATTAA	TGCTAATATT	AATAAAGTTA	93540
ATTGTCTTTA	AGGATTTAAC	CGTGGTAAGA	GGAATTTATA	CAGCTGCCAG	CGGAATGATG	93600
GCAGAAAGGC	GCAAGCTTGA	TACCGTGTCA	AATAATTTGG	CAAACATAGA	TCTTATTGGA	93660
TACAAAAAAG	ATTTGTCTAT	TCAAAAAGCA	TTTCCAGAAA	TGCTAATAAG	AAGACTAAAT	93720
GATGATGGTC	TTTATAAATT	TCCCAAAGGA	CATCTTGAAA	CAGCTCCGGT	TGTGGGCAAA	93780
ATAGGAACAG	GGGTGAAGA	AAATGAGATA	TACACAGTAT	TTGAACAGGG	CCCATTAAAA	93840
ACTACTGGCA	ATCCATTAGA	TTTAGCAGTC	ACCGATCAAG	GATTTTTTCGT	AATACAAACT	93900
TCAGATGGAG	AAAGATATAC	AAGAAACGGT	TCTTTTACTA	TTGGAAAAGA	AGGAATCCTT	93960
GTTACAAAAA	GCGGATTTCC	CGTTC TAGGA	GAAAAAGGAT	ACATATATCT	TAAGAAAAAT	94020
AATTTTAAAA	TAACACCTCA	AGGACAAGTC	TTTCACAATT	CAAACCTTGA	ATCAGACCCC	94080
AAAAGACTTG	TTAGCGAGTA	TGAAAATTCT	TGGGAAAATT	ATGAGCTGCT	TGATACCATT	94140
AGAATTGTAA	ATTTTGAAAA	TCCCAGATTT	CTCAAAAAAC	AGGGAAATTC	TTTATGGATC	94200
GATACAAAAA	CATCTGGCAA	AGCACAAGAA	ATTGATATAT	CATTAAGGCC	TAAAATAGAA	94260
ACAGAAACAC	TTGAGGCTTC	CAATGTTAAT	GCTGTTAAAG	AAATGGTTTT	AATGATTGAA	94320
ATTAACAGAG	CTTATGAAGC	TAATCAAAAA	ACAATACAGA	CTGAAGATAG	TCTATTGGGA	94380
AAATTAATAA	ATGAAATTGG	AAAAATATTAA	GGAGCATGTT	TTATGATGAG	AGCATTATGG	94440
ACAGCAGCAA	GTGGAATGAC	TGCACAACAA	TACAATGTAG	ATACAATTGC	CAATAACCTT	94500
TCAAATGTAA	ATACTACAGG	ATTTAAAAAA	ATAAGAGCAG	AATTTGAGGA	TCTAATTTAT	94560
CAAACCCATA	ACAGAGCAGG	AACCCCTGCA	ACTGAAAATA	CTTTAAGACC	ACTTGGAAT	94620
CAAGTTGGTC	ACGGAACAAA	AATTGCTGCC	ACCCAGAGAA	TATTTGAACA	AGGAAAAATG	94680
CAATCCACAA	ATTTACTCAC	TGACGTTGCC	ATTGAAGGAG	ATGGATTTTA	CAAATTTCTT	94740

CTACCTGATG	GAACCTTATGC	ATACTAGA	GATGGGTCAT	TTAAAATCGA	TCTAATCGA	94800
GAGCTTGTA	CAAGCCAAGG	ATACAAAGTA	TTGCCTAATA	TACTCTTCCC	AGAAGAATAT	94860
ATCCAAAAC	CAATTACAAT	ATCTGAAGAG	GGAATAGTAT	CGGTAAAAAT	TGATACCAGC	94920
AACGAACCA	TAGAGCTTGG	GCAAATTGAA	ATATCAAGAT	TTATCAATCC	TGCAGGACTA	94980
AGTGCCATTG	GAAGCAATTT	ATTTAAAGAA	ACAGCTGGAT	CAGGCCAAGA	AATAGCAGGA	95040
ATACCAGGAA	GTGAAGGCAT	GGGAAGACTA	AGGCAAGGCA	TACTTGAAAT	GTCAAATGTA	95100
TCTATTGCTG	AAGAAATGGT	AACAATGATA	GTAGCTCAAA	GGGCTTATGA	AATAAACTCA	95160
AAAGCTATTC	AAACTTCTGA	CAATATGTTA	GGAATTGCAA	ATAACTTAAA	AAGGCAATAA	95220
AATAAAAAAA	AGATTATTTA	TTTTTATTTT	ATTTTTCACA	ACAAGCTCAA	TTATAAGAGC	95280
TTCTCATGAT	TTATGTTTCA	ACATTGCGCC	TAGTAAAACA	TATTTCTTTT	CAAAGAAGTA	95340
TTCAAAAATA	TGTAACAATC	AAAGCTTATC	AAAAATATAT	ATCCCCCAC	ATTTAACAAA	95400
AAAATCAATA	ATTTTGTAAA	TGATTTATTA	CATTACAAAA	AATTTATCAA	ATGAAAATAT	95460
CTATATACTT	CAATTTAACT	TTGATGAATC	TGAAATAAAC	ATAGAAGATA	AATTTTTCAA	95520
AAAAGTAAAA	TTTAAGGTAA	AAAGCAACAA	TTCATACAAA	AATATTCCAA	TTGAAAAAAC	95580
TCTTGTTTAT	TATGCAAAAA	ACTTTGAAAG	CTACAAAAGA	CACAATTACA	TCAATATGTA	95640
CATTGATGTA	ATCGAGCCAA	TTGTATTTGC	AAAAGAAAAT	CTAAAAAAA	ATGAAATCCT	95700
TAATGAGTAC	AATACATACT	TTAAATACAA	AATTAACACA	ACAAGAATAA	ATGATGTTTT	95760
AAGTCTAAAT	GAATTAAACA	ATAGCAAATA	CAAAGTTATA	CGCAACACAA	TCAAAAATGA	95820
AGAGATAAGA	TTAAATAAGG	TGCAAAAAGA	ATAATACCTA	ATTTTATCTT	CCTTTTCTAA	95880
AAATTATTAT	TTTAATCTCC	CTTAATGCAG	CTAATATTTA	ACAAATCAAG	GATTAATTAG	95940
TAATTTAACG	AAAAAAGTTT	CATTAATTGC	AATAATTGAT	ATAAAATAAT	AGATATTAAA	96000
GAAATACAAT	AAATAAGGTA	AAGAATGAAC	AAACTAATGT	TGATGTTAAT	TACATTTGCA	96060
ACGAGTCTAT	TAGCCCAAAC	AAACAAAGCT	TCAACAGGAC	TAAAAACAGA	TCAATCATTT	96120
AACAATAGCC	TATCTGAAAG	CGTAAATTA	AAAGAAATTG	CGGATATTTA	TCCCACAAAT	96180
ACAAATTTTT	TAACAGGTAT	TGGAATAGTA	GCGGGACTTG	CTGGAAAAGG	AGACTCTATA	96240
AAACAAAAAG	ACCTTATAAT	TAAAATTTTA	GAAGAAAACA	ATATAATAAA	TGAAATAGGC	96300
TCTAATAACA	TAGAAAGTAA	AAATATTGCA	CTAGTAAATG	TCAGTCTCCA	AGTAAAAGGT	96360
AATACAATCA	AAGGTTCAAA	ACATAAAGCT	TGCGTTGCAT	CAATACTGGA	CTCAAAAGAT	96420
TTAACAAATG	GAATACTTTT	AAAAACAAAT	CTTAAAAATA	AAGAGGGGGA	AATAATAGCA	96480
ATTGCATCAG	GAATTACACA	GCCCAATAAT	AAATTAAAAG	GATCTGGATA	TACTATAGAT	96540

AGTGTAATAA	TAAATGAGAA	TCAAAATATT	AACCACAGTT	ATAATATAAT	TCTTAAAAAA	96600
GGAAATTATA	CATTAATAAA	TAGAATTCAT	AAAATATTAA	CCTCTAAAAA	AATCAACAAC	96660
AAAATTAAAT	CAGACAGCAC	AATAGAAATA	GAAGCAAAAA	ACATAAGCCT	ATTAGAAGAG	96720
ATTGAAAATA	TTAAAATAGA	AACCAACCCC	AAGATATTAA	TAGACAAAAA	AAATGGTATT	96780
ATTTTAGCAA	GTGAAAATGC	AAAAATAGGA	ACTTTTACAT	TTTCCATTGA	AAAAGACAAT	96840
CAAAACATTT	TTTTAAGTAA	AAATAACAAA	ACAACAATTC	AAGTAAACTC	AATGAAATTA	96900
AATGAATTTA	TATTAAAAAA	TTCCAACAAT	CTTAGCAATA	AAGAATTAAT	TCAAATAATT	96960
CAAGCTGCGC	AAAAAATTAA	TAAATTAAAT	GGGGAACCTA	TCTTGGAGGA	AATTGATGGA	97020
AACCAAAATT	AATTCACAAA	ATCTAAAATT	TAAAAATCAA	ATAAATAATT	TTAAAAATTC	97080
TGTAGAAATA	AAAAAATCCT	TTCAAAAAAA	CGAAGATCTT	CGAAAAGCTT	CTTTAGAAAT	97140
TGAAGCTATG	TTTATCAAGC	AAATGCCTTG	AAGCATGAAA	AAAACCTCTT	ACAAAGATCA	97200
AAATTTGCTA	AACGGAGGCC	AAGTAGAAGA	AATTTTGTGA	GATATGCTTT	GCGAACAAAG	97260
AGCAAAACAA	ATGGCACAAAG	CTCAAAGCTT	TGGCCTTGCC	GATTTAATTT	ACAATCAATT	97320
ACAAAAAAGT	AAATAATTCA	AAAAATACTC	CCCCTAAACT	CAAATTATA	TCCTATTTAG	97380
TTTAAAACCA	TTTTTTAAAT	AAATTGGCAC	AGTTTTTGCA	TGGAAATTAA	GTAGTAAAAA	97440
CTTAATCACA	ATATTCAAGA	AAGGGGAGAA	AATATAATAA	CTATGAACAT	ATTTAGTAAT	97500
GAGGATTTAA	ACATATATTT	AAAATCAGTA	AGAGAACACA	AGCTAATTAC	TCACGAAGAA	97560
GAAATCAAAC	TTGCAGGACA	AATACAAAGA	GGCAATGCAA	AAGCAAAAAA	CAAGATGATA	97620
AATGCAAACT	TGCGACTTGT	TTTAAAAATA	ATAAAAAGAT	ATGCGGGTAA	AGGGTTAAAA	97680
ATTGAAGACT	TAATTCAAGA	AGGCAACTTG	GGATTAATAA	GAGCTGCTGA	AAAATATGAC	97740
CCGAATAAAA	ATACCAAATT	TTCAACTTAT	GCATCATTTT	GGATTAAGCA	ATCACTACAA	97800
AGAGCATTA	ACACTAAAAC	CAGATTGGTA	AAAGTCCCAT	ACAGAAAAGA	AAATCTAATA	97860
CTACAAATA	ATAAATATTT	AACAGAAGAA	GAAAAATCGC	CAAAAAAGA	AGAAATAATG	97920
AAAAGATTCA	ACCTATCTCC	TGCTCAGTAT	ATAAAAATTA	TTCCCTATCT	TGAAAAAGAA	97980
TATTCTCTGG	ACAAAGAAAT	AGAGGGATCT	GAAAATTCAA	CACTCTTGAA	TCTATACGAG	98040
GATAATTCTT	TTAACCCCTGA	AATTACCCTT	GAACAAGATT	CAACTCTAAA	ACATTTGAAT	98100
TATATACTTG	AAACAAAATT	AAATGAAAAG	GAAAGATACA	TAATTAAAAA	AAGATATAAC	98160
CTGGACAATA	GTCCCAAAAA	AAGCACCTTA	AAAGATATTT	CAACAGAACT	TGGAATATCA	98220
TCAGAAACTG	TAAGACAGAT	TGAAAAAAGA	GTTCTTAAAA	AATTAAAAGA	AGAAATAAAT	98280

TAACATTGAC	ATTCATGACA	TGTTCTGGTC	TACTTGTAAG	TCAGTGGTCA	TGAATGTTTG	98340
TATTTTATAT	TAAAAAAAC	AAGTTTATTA	TTGTAATTTT	TATTATAATT	TCTATTGTTA	98400
TTGCAATAAC	TCAGGCATTT	GCAAGTTTTT	TATATTTTAA	TGACAATTCA	AAAATTGCAA	98460
ATGCCCCACT	TAAAAATAGG	TTTGAAAAAA	CACAAAAAGA	AAGCTTAATA	ATAAAAACA	98520
ACAACGAGGA	TAAAAAAGCC	AAAAGCAAAC	CTAAGTTTTA	CTTAATCATT	GACGACGTGG	98580
GCTATGATGA	ATTTATGTTA	GAACAATTTA	TAAAACTTAA	TCTTAAAATA	ACTTATGCTA	98640
TTATTCCATT	TTTACCAAAA	TCAATGAGTT	TATACAAAAA	ACTAAAAAAT	GCTAACAAAA	98700
CAGTAATAAT	ACATTTCCCA	ATGCAATCAA	AACATAGAAA	TTCAATAGAA	AAATTTTATA	98760
TAAACATAAA	AGATAAAAAA	GAAGAAATAC	ACAAAAAAAT	CGAAAAAGCA	TTTAAAAAGT	98820
ATCCTGATGC	AAAAATAATG	AATAACCATA	TGGGAAGTTT	AATCACTTCA	AATAAAGATT	98880
TGATGAAAAT	CATTTTAGAA	AAGCTTAAAG	AGATTGACAG	ATATTTTTTC	GACAGCGTAA	98940
CTATTGCAGG	AAGCGTACCA	GAAATAATAG	GCAAAGAAAT	TGGAGTTAAA	GTAGAAAAAA	99000
GAGACGTATT	TCTTGATAGC	AAAGACACAG	AAGAGTCCGT	AACAAAGGAG	CTTGAAAAAG	99060
CAAAAAATAT	TGCTAGAAAA	AATGGAATGG	TAAAAGTAAT	AGGACACATT	TGGTCTAAAA	99120
ATACGCTAAA	AGTCCTTAAA	AAAGAAGGAC	CTGATTTAAA	CCAGGAATTC	GAATTCGACA	99180
ACTTATTAAA	TCTTTACGAG	GAAACAATCA	GATGAAAGTG	CTTGGAATAG	AAACCTCTTG	99240
TGACGACTGT	TGCGTAGCTG	TAGTAGAAAA	TGGAATTCAT	ATTTTAAGCA	ATATAAAATT	99300
AAATCAAACC	GAACACAAAA	AATATTACGG	CATAGTGCCT	GAGATTGCCT	CAAGACTTCA	99360
TACGGAAGCT	ATTATGTCTG	TTTGTATAAA	AGCACTAAAA	AAGGCAAATA	CTAAAATATC	99420
TGAAATTGAC	TTAATAGCTG	TAACATCTAG	ACCTGGACTT	ATTGGATCTT	TAATAGTTGG	99480
ATTAAACTTT	GCCAAAGGTC	TAGCAATTTT	ATTAAAAAAG	CCCATTATTT	GCATTGATCA	99540
CATCTTGGGT	CATCTTTACG	CCCCTTTAAT	GCACTCAAAA	ATAGAATATC	CATTTATATC	99600
ATTATTATTA	AGTGGTGGAC	ATACATTGAT	TGCTAAACAA	AAAAATTTCG	ATGATGTTGA	99660
AATACTTGGA	AGAACTCTAG	ATGATGCTTG	TGGAGAGGCT	TTTGATAAAG	TGGCAAAACA	99720
TTATGATATG	GGATTTCCGG	GAGGTCCAAA	CATCGAACAA	ATATCTAAAA	ATGGAGATGA	99780
AAATACATTT	CAATTTCCAG	TTACCACCTT	TAAAAAATAA	GAAAACTGGT	ATGATTTTTTC	99840
ATACTCTGGA	CTAAAAACAG	CTTGCATACA	CCAACTCGAA	AAATTCAAAA	GCAAAGATAA	99900
CCCAACAAAC	AAAAATAATA	TAGCTGCAAG	CTTCCAAAAA	GCTGCCTTTG	AAAATCTAAT	99960
CACCCCACTA	AAAAGGGCAA	TAAAAGATAC	TCAAATCAAC	AAATTGGTAA	TAGCAGGAGG	100020
TGTTGCAAGC	AATTTATATT	TAAGAGAAAA	AATAGATAAG	CTTAAAAATAC	AAACTTACTA	100080

CCCTCCTCTT	GACCTTTGCA	CAGACAATGG	AGCAATGATT	GCGGGACTTG	GATTTAATAT	100140
GTATTTAAAA	TATGGAGAAA	GTCCAATTGA	AATTGATGCA	AATTCAAGAA	TAGAAAATTA	100200
TAAAAACCAG	TATAGGGGGA	AAAATAATGA	AAAGAATTTT	AGCAATGCAT	GATATTTCAA	100260
GCATGGGAAG	AACATCTCTT	ACAATATGCA	TACCAGTAAT	ATCTTCGTTT	AATATGCAAG	100320
TTTGTCTTTT	TGTGACAGCT	GTCCTTTCTG	CTTCCACAGC	TTATAAAAAA	TTTGAAATAG	100380
TGGATTTAAC	CGATCATTTA	GAAAAATTTA	TCAATATATG	GAAAGAACAA	AATGAGCACT	100440
TTGACATACT	CTATACCGGA	TTTCTGGGAA	GCGAAAAACA	ACAAATAACA	ATAGAGAAAA	100500
TAATTAAATT	AATAAAATTT	GAAAAATTG	TAATTGATCC	TGTGTTTGCT	GACGATGGAG	100560
AAATTTACCC	TATATTTGAT	AATAAAATAA	TTAGTGGATT	TAGAAAAATC	ATAAAGTACG	100620
CAACATAAT	AACACCCAAT	ATCACAGAAC	TTGAAATGCT	AAGCAAAAGC	TCAAACTTA	100680
ACAACAAAGA	TGATATCATA	AAAGCAATAT	TAAATCTTGA	TACAAAAGCG	ACGGTAGTTG	100740
TTACAAGCGT	TAAAAGGGGA	AATCTCTTGG	GAAACATTTG	CTACAATCCT	AAAAACAAAG	100800
AATACTCGGA	GTTTTTTTTT	GAAGGATTAG	AACAAAATTT	CAGTGGAACA	GGAGATTTAT	100860
TTACCAGCTT	ACTTATAGGA	TATTTGGAAA	AATTTGAAAC	AGAGCAAGCC	TTAGAAAAAA	100920
CAACAAAGGC	TATTCACCTA	ATAATAAAG	AGTCAATTAA	AGAAAATGTT	TCAAAAAAAG	100980
AAGGGGTCCG	AATTGAAAAAT	TTCTTAAAAA	ATACATTTTG	AATTTAAATT	CCATTAAATT	101040
CAATTTTTTAA	GATTGAATCA	ATTCTTTGGT	ACAAAGGAAA	TACTGATATT	GCAATATATT	101100
ATTAAAATAA	AATGTGAAAA	AATTTATTAC	AAAGTAAATG	CTTTATTGTT	TTCATGAGTA	101160
AATAAAAATA	TGTCAAATAA	AAAAATAATA	TTTTTTTACAG	GGGGAGGAAC	TGGGGGTCAC	101220
GATTTTCCAG	GAATTTCCAT	CATACAAAAA	TTAAAAGAAT	TTGATAATGA	AATTGAATTT	101280
TTTTGGATAG	GTAAAAAAA	TTCTATAGAA	GAAAAACTAA	TAAAAGAACA	AGATAATATT	101340
AAATTTATTT	CGATTCCATG	CGGAAAACTT	AGACGCTATT	TTTCTTTTAA	AAATTTTACT	101400
GACTTTTTTCA	AAGTAATACT	TGGAATAATA	AAAAGCTTTT	ACGTTTTTAA	AAAATATAAA	101460
CCTCAGCTTA	TTTACGCAAC	CGGAGGATTT	GTTTCAACTC	CTGCAATTAT	TGCATCCAGC	101520
TTGCTAAAAA	TAAAAAGCAT	AACCCATGAA	ATGGATCTAG	ATCCCGGACT	TGCAACAAAA	101580
ATTAECTCTA	AATTGCAAAA	TAACATACAC	ATAAGCTTTA	AAGAAAGTGA	AAAATACTTC	101640
AAAAATTACA	AAAACATTAT	TTACACAGGA	TCTCCTATAA	GAAGAGAATT	TTTAAATCCA	101700
GATCCCAAAA	TAATCAAACA	ATTGACACAA	AACACTAACA	AACCAATTAT	TAGCATACTT	101760
GGGGGATCTC	TTGGCGCTAA	TGCTTTAAAC	AACCTTGCAC	TCTGCATTAA	AAAGGATGCT	101820

GAAATCTACT	TCATCCATCA	ATCGGGGAAA	AATTTAAATG	ACCTAAGCGA	AAAGAATTAC	101880
CTTAGAAGGC	AATTTTTTAA	CGCAGAAGAA	ATGGCAAGTA	TAGTTAAATT	TTCTAATCTA	101940
ATAATAAGCA	GAGCCGGAGC	TGGAGCAATA	AAGGAATTTG	CAAATGCTGG	TGCATGTGCA	102000
ATTTTGATTC	CATTTAAAAA	AGGCTCTAGA	GGAGATCAAA	TTAAAAATGC	AAAATTACTA	102060
ACAAATCAAA	ATGCCTGCAT	TTATATAGAT	GAAGATGAAA	TTTTAAATAT	AAATATTTTA	102120
AAAATTATAA	AAAAAACTTT	AAAAGATAGA	GAAAAAATCA	ACTCTCTCAA	AGAAAAATATC	102180
AAAAAATTCA	ATAATAAGCA	TTCTTCAACT	TTAATAGCCA	AATTGCTAAT	AAAAGATATT	102240
AAGGAGACAA	AATCTAAATG	ATAATAAACG	ATCCTGTAAA	AATAACTGGA	ATAGTAGACA	102300
TATTAATAAT	AATAATTTTT	ACATCTTTGG	GATTTAGAGG	ATTTTAAAGA	GGATTTATTA	102360
AAGAAATTAG	CGGATTTGCT	GAAGTTTTTG	TTTTAATCCT	ACTGCTTTAC	AAAAAACTG	102420
AAGAATTTAG	AAGGTTTGTT	GAACCTATTA	TTGAGCTATC	CTACATTCAA	GCACTACTTG	102480
TATTTTTTTT	GCTTATACAT	ATAGGATTTT	TAATACTACA	ATCCCTAATA	GAATCAATAA	102540
TAAGTCAACT	TAAATTGCTA	TTCTTCAATA	GAATACTAGG	CTTAGTGCTT	GGCCTACTTG	102600
AAGCTTTTGG	AATAATTGCA	ATCGTGGTTT	ACATAATACA	CTCACAACAA	ATATTTAAAC	102660
CTGAATATTT	CCTAAAAGAA	AGCAAACCTAC	TTGATTATTT	AAATCCTGGA	ATAAACTATC	102720
TCTTTAAAT	TTCAAAAACA	AAATAAGGGC	CAGCAATGAC	AATGCTTCCA	AAAATTGCAA	102780
AAGAGATAAT	AAACGAATAT	GATCAAAAAA	TACTGCCCAA	TGCAATTCTT	TTACTAGGAG	102840
AAAAATTTTC	TTCAAAAAG	ATTAGCGCAA	TTGAGCTTGC	AAAAAAAATA	TTAAACGGAA	102900
AAAACCTAAC	AAACCCTAAT	TTGCTCATTT	TCTCAAATCT	TGACACAGTA	GAAGCAAAAG	102960
CACATCTTTC	TACAAATTCG	CAAAAGATAG	CAAATAAATA	CCTAGAATAT	ATTAAAACTG	103020
TAATTTTTTAC	CAAATGTTAT	TTCAAGCAATG	AAAAAAATTT	AAAAAAAATA	GAAAAAAATA	103080
TCAACTACAT	TAATTCTGTT	TATTATGAAA	AAGAATACAA	TGAAAACATA	AAAAATGAGC	103140
TTATAAAAAA	TATAGAAAAT	ATAAACAAAG	AATTAAATCA	TAGCATTACT	GTTTATGATG	103200
TAAAAAAAT	TCAAACCTGG	ATTTTTTCTG	AAAAAGAAAA	ACCAAAGGTA	ATCTACATAA	103260
ACGAAATCGA	AAATTTATCA	TTTAATGTCC	ATAACTCACT	TTTAAAAATA	TTGGAAGAGC	103320
CTCCCTCAAA	TATTTACTTT	ATCTTGCCAG	CAAGAAATAA	AAACAAAATA	CCAAAAACAA	103380
TACTTTCAAG	ACTTAGAGTC	TACAATTTTCG	CAAACTAGA	CAGAAGCTTA	GAAATTCAAA	103440
GATTTAAAGA	AAGCTTTCTA	ATAAATAAAG	ATATAACAAT	TGAAGAGTAT	TTGCCTCAT	103500
TTTACAAAGA	AGAAAGCAAA	AAAATAAAAA	AAGAATTGGC	AAAAATTCTA	AATATAATAA	103560
AAGAAAAAAA	ATCCATATTT	AATCTTGAAG	AAGTCGACTT	TATAAAAGAT	GAGCAAAGCT	103620

TTAAAATATT	TTTAAACGAA	CTTACAATTA	ACATTAGAAA	AGATTTTTTTA	GAAAACAAAA	103680
TAGATATTAA	TCAATATCTA	AAGTACACAG	AGCATTTGAA	AAATATTTAC	AAATATCGCC	103740
CCTATAATCA	AAATAAAAAA	TTAATAATAG	AAAACCTAAT	GCTAAATTAT	GAGGAGATAT	103800
GAATAATTTT	TTCAAAAAAG	CTTTAACAAA	GCTAAACAAA	TTATCTAACG	AACAAAAAAC	103860
TAAATTTATT	GAACAAATTT	ACAAAAAAAT	AGAAATATAT	GACGGAATAT	TTGCATCAAT	103920
TAATGAAGGA	ATCATTGTAC	TTGACAAACA	AAACAATATA	ATCTATGCAA	ACAAGATTTT	103980
ATACCAAATT	TTAGCTTTAA	CATCTAAATC	AAAAATAGAA	ATTCTTGATG	ACATTCAAAT	104040
TCCAAACTTA	ATAAATTTAA	TAAAAGAACT	AGTTAGAACA	GAAGATAAAA	TAATAGGATT	104100
AGAAGTTCCA	ATCTCAAACG	GCATATATAT	TAAAATCTCA	TTTATGCCTT	ATGTAAAAGA	104160
AAAAAACTT	GAAGGCAACA	TTATTTTAAT	CGAAGACATT	AAAGAGAAAA	AAAAGAAAGA	104220
GGAAC TATTT	AGAAGAGTTG	AGGCTTTGGC	CTCTTTTACA	AGGCATGCAA	GAAATATTGC	104280
CCATGAAATC	AAAAACCCAC	TTGGAGCAAT	CGATATAAAT	TTACAACTGC	TAAAAAAGGA	104340
AATTGAAAAA	CAAAAAATGA	AAAATGGTAA	AGCTGAAAAT	TATTTTAAAG	TAATAAAAGA	104400
AGAAATAAAC	AGAGTAGATA	AAATAGTAAC	AGAATTTTTA	CTAACTGTCA	GACCAATAAA	104460
AATTAACTTA	CAAGAAAAAG	ATATTAAACA	AGTAATAGGC	AGCGTATGTG	AATTGTTAAA	104520
TCCTGGATTA	GAAAATAAAC	ACATAAAACT	ATTGCTTAAT	TTAAACAAAA	TAAGCAATAT	104580
TCTCATTGAT	GAAAACTAT	TAAAACAAGT	TATTATAAAC	ATCGTTAAAA	ACGCAGAAGA	104640
AGCACTGCTT	GAAACAAAAA	AAGAAATAAA	AAAAATAGAA	ATTTTCTCT	TCGAAAAAGA	104700
CAATAAAATA	CATATCAACA	TAAAAGATAA	CGGAAACGGA	ATAAAAGATG	GGGTAAAAGA	104760
GGAAATATTT	AAGCCTCAAT	TTAGCACAAA	AGAAAAAGGA	AGTGGAAATAG	GACTTACTAT	104820
TTCTTATAAA	ATAATAAAAG	AGCTTGGAGG	TGAAATTTTT	GTGGAAAGCA	AAGAGGGCAA	104880
AGGCACTATT	TTTACAATTA	CGCTGCCTAA	ACTAAATAAA	AAAAATATTT	TAATTGAAGG	104940
GTATTGAAAA	TGAGCAAAAT	ACTTGTAGCT	GATGATGAAA	AGAATATTAG	AGAAGGAATT	105000
GCTACTTATC	TTGAGGATGA	AGGATATTTT	GTTTTCACTG	CTAGTGACGG	AGAAGAAGCT	105060
CTTGAAACAA	TTGAAAATGA	AAATCTTGAT	GTAATAATAT	CTGACCTGAG	AATGCCCCAG	105120
ATATCTGGAG	AAAAATTGCT	CAAAATAGTT	AAAGAAAAAA	ACTTGGAAT	ACCTTTTATT	105180
ATTCTAACAG	CCCACGGAAC	AGTTGATTCT	GCTGTAGATG	CCATGAGAGA	GGGTGCTTAT	105240
GATTTTTTAA	CAAAGCCCTT	AGACCTTGAA	AGACTTTTGC	TAATAATAAA	AAGATCACTA	105300
AATAAAAAAG	AAAATAACGA	TAATGAAAAT	GCTAATTTAG	AAAATATACT	AATAAGAAAA	105360

GATCTAAAAT	ACTATGAAAA	ATCATGCGGA	AAATCCCTAT	TAATGCAAAA	ATTTTTTGAA	105420
CTTGTAATAA	AAATAGCAAA	ATCAAATGCA	TCTATTCTTA	TAACGGGCGA	AAGCGGTGTT	105480
GGTAAAGAAA	TAATAGCAGA	TGCTATTTTT	GATCTTTCAA	ATAGAAATGA	CAAACCATT	105540
ATAAAAGTAA	ATTGCGCAGC	ACTTTCTGAA	AGCATTCTTG	AAAGTGAAC	TTTTGGCCAT	105600
GAAAAAGGAG	CATTCACTGG	AGCAATTTCC	AAAAAAAAG	GCAGATTGGA	ACTTGCAAAC	105660
AAAGGCACAA	TTTTTCTTGA	TGAGATAGCA	GAAATTTTCA	CTGAAATTCA	AGTCAAGCTT	105720
TTAAGAGTAC	TGCAAAACAA	AACTTTTGAA	CGTGTGGGG	GAGAAGCTAC	AATTAAAGTT	105780
GATATCAGGC	TTCTGGCTGC	AACAAACAAA	AACATTGAAG	AGGAAATTAA	AAAGGAAAAA	105840
TTTAGAGAAG	ATTTATTTTA	TAGATTAAAT	ATCATTAATA	TAAACATACC	GCCTTTAAGA	105900
GAAAGAAAAG	ATGATATATC	TTATTTAACA	AACATACTAA	TAAAAGACGT	CGCAAAGGAA	105960
AACAATAGAG	AAGAAAAAAC	TCTTTCTAAT	GATGCAATGA	AAGCTCTCTA	TTATTACGAT	106020
TGGCCAGGAA	ATATTAGAGA	ATTAAAAAAT	GTGCTTGAAA	GTGCATTAAT	ATTATCAAAA	106080
GGCAAACAAA	TCACTAAAGA	AGATTTGCCA	GCAAAAATCA	AAAATAATGA	AAATCTTATA	106140
TTTAAATATA	CACTACCAAT	AGGAATTAGC	CTAAAAGAAG	CTGAAAAAGA	AATAATAAAA	106200
CAAACACTTT	TTCAATCCAA	AAACAACAAA	AGCAAATGCG	CCGAAATACT	AAAAATAGGA	106260
AGAAAACTT	TACACAATAA	AATAATCGAA	TATAATATTG	ATTAATAGGA	TTATTTTAA	106320
ATTATTAAAT	TATAATGGGT	ACAAAAAAT	AATACTGCTT	TAAATTCAT	GTATATTTTT	106380
GAAACCAAAA	AATTTTTTAA	TGCCAATAAT	TATATTAAAA	TGAAACACTT	TCTTTTAAAA	106440
TCATGGCGCA	AAAGTGTAAG	AATATTTTTA	TCAAACAAAT	AATTATACAC	CATTATTTGT	106500
TAATAATCAA	TACAATTTGA	TAATTTAATA	TATTTAGCTG	GCTACAGAGC	CTGACCTTAC	106560
TTTAAAACT	TTAAAGGGT	AATAGGAATA	TTTTTTTTTA	ATATTTCAAA	GTGCAAATGA	106620
GGACCAGTTG	CGCGACCCGT	TTGCCCAACC	ATTCCAAGAA	ATTCTCCCGA	TTTAACAAAA	106680
TCACCTATCT	TTACAGAATA	TAAATTTAAA	TGCCCATAAA	GAGATTTAAT	ATTATTTTTG	106740
TGACCAACCA	CAACAAAATT	CCCATAAAGA	TCATTGTATC	CAGCTTCAAT	AACTATTCCA	106800
GAAGAAGAAG	ATACACTTCA	GCATTCATTG	GAGCTGCAAG	ATCTATTCCT	GTATGGAAAC	106860
TTTTGTTGCC	AGTGAAAGGG	TCATTTCTAA	ATCCAAAATC	AGAACTAACA	ATAAATTTTT	106920
TTAAAGGAAA	AATAAAATTG	GCATTTAAGA	AAAAAAGCAA	TTCTGTGCCT	GAAAAAAGTC	106980
CAAAATCTGG	ATTCTTAACA	AAATCAAAAA	AATAAAATTC	ATAAACTCTG	TCGTTCCCTT	107040
TAATTTTTAC	CTTTTCAGCT	TTAGCAAGAT	CCCTTGTTGC	TAAAAGCAAA	TTATTAAATC	107100
TATAATCTTT	ACTATCAAAA	ACAAAAACTC	CTTTTTTACT	GGAATAAGA	ATCTCTTGCC	107160

CAACACTCAC	AGCAGGAGAA	TCTAATAAAT	TAATAGTAGC	AATGCCGGAC	TGCCATCCAT	107220
TTATTTTATT	GGCAATTTTA	AAAAAAGTAT	CCCCTTTTTT	AACTTTATAT	GAGTAAAAAA	107280
ACAGAGGAAT	ATGTTGTTTT	TTGTTATATT	TTAAAACTTT	AATTTTAAGA	TCAGAAAAAA	107340
CAGGATCTTG	CCTTGAGAAA	TTTTTTATTT	CTGGATAAGA	AAAAACATAA	ATTATTTTTA	107400
AAAAAAGAA	ACCTGCATTA	AATAATAAAA	AAATTTTACT	CATACTATAA	ATTCTTTAAC	107460
GATTAATTAA	TCAAATATAA	TAAAAAACAT	AAAAAAAATA	AATCCTATTT	GGACTTGCAA	107520
ATAATAACAA	ACTGTAATAA	ACTGTCTCTC	AACATGGAGC	TAAACGAATA	CCAAGAAAAA	107580
GCAAAAAAAA	CTGCTAAATA	CAAAAATAAA	AAAGAAGAAT	TAATTTTAAC	AACACTTGGT	107640
CTTGCTGGTG	AACTGGAGA	AGTTGTTGAA	AAAATAAAAA	AATTGGGAAG	AGATAAAAAAT	107700
TACATTATTG	ATGATGAGTA	TTAATATCA	ATTAAAAAAG	AGCTTGGGGA	CGTATTATGG	107760
TACTTGTCAA	GTTTAAGCAA	TAATTTAGGC	ATTACGCTTG	AAGATGTTGC	CCTCACAAAC	107820
CTAAAAAATA	TACAAAAACG	ACATGAAAAAT	GGAACAATAA	ATGGCGAAGG	CGATGACAGA	107880
TAAGGCATTT	AAATTTAAAA	TACTCAAAGT	TTAAAAATAA	AAATGCTTAA	TATTTATATC	107940
AAGGGAATTT	TACTTGGAAT	TGCAAACATA	ATCCCAGGGG	TTTCTGGGGG	AACGCTGGCT	108000
TTAATATTAA	AAATTTATTA	CAAAATAATA	AACTCCATCT	CAGAAATCTT	AAAGCTCACA	108060
GAAATTAAAA	AAAATTTAAT	GTTTTTAACT	ATTTTGCCAA	CAGGAATGTT	AACCTCAATA	108120
TTATTAACTG	CAAAAATATT	TAAAACTTAT	GCTTTTGACA	ATGGAATAAT	AGAAGCACTG	108180
CTAATAGTAT	TTTTCATAGG	ATTAGCATTT	GGAAATATAC	TAACACTAAA	AACAGAAATA	108240
TCTATAAAAAG	AAATAAATAG	TAATACAAAA	ATATTAAATA	ATTTATTGTT	TTTCATTGGT	108300
ATGAGCATTA	TTGTACTCTT	CTTAATACTC	AAAGAATCTA	ATATACAATT	GCAAAGTACA	108360
ATACCTAAAG	ACAAAACTC	AATAAAATAT	TACTTATTAT	TGATATCCTC	TGGAACAATA	108420
AGCGGAGCAT	CAATGATCTT	ACCGGGAATC	TCAGGATCTG	CAATGCTTTT	ACTGCTTGGC	108480
TTTTATAAAG	AAATAATACT	TATTGTGTCT	GAATTTAACA	TTATTCTTAT	TACAATATTT	108540
GCAGCTGCTG	CAACAATGGG	AATAATTACA	TCAATATTAA	TAATAAAGAA	AATAATAGAT	108600
AAGCACTTAA	ATAATTTTAT	TTATTTATCA	AAAGGCTTAA	TTTTTGATC	AATTCTACAA	108660
ATGATATTAA	TTGTATTAAA	ATTGAACTTT	AAAATCGGCT	TTACATCTTT	TACATCTCTG	108720
GGAACATCAT	TCATACTGGG	AATCTTTATA	AACAAAAAAT	TGGCTGAGAA	ATATAAATAA	108780
AAAATTTAAA	AATACCGAAG	ACCGGACTTG	AACCGGTACG	AGCTTCCTCC	TCAGGATTTT	108840
AAGTCCTGTG	TGTCTACCAA	TTCCACCACT	TCGGCATAGA	ATAATATAAT	AAATAATAAT	108900

ATATTGCAAT	GTCAAGTTAA	TAAGAAAAT	AAATATAAAA	ACTCAATTGA	TAACATTTTT	108960
TTGAGGAAAA	TTCTCAACAA	CATAATTGCT	ATGCAAATAA	TCTGCAATTA	CAAACATAAA	109020
TGAATATAAA	GAAACCCTAG	CAAAAAAATT	AGAAAAAACT	CTAAAATCAA	TTTTATTTGC	109080
TAACATTGAA	GTGTAGTTTG	AATCATTTAT	TAAGCTAAAA	AATCCATTTT	TGATTAATAA	109140
ATCAAAAATA	ACAAATTCAG	ATTCATCTAA	AATAAACATC	AATTTATCCA	TTTGGGCTTT	109200
TAAATATAAA	ATTTGTTTAC	TAATTTGCAA	ATTGGCAATA	CTTTTAATAA	AATCCTTGCT	109260
TAAATTATAA	ATAACAATAC	CATAGTGATT	ATTCTCAAAT	TTAAAATCCA	CAAGCAATCC	109320
TAATTTTAAA	TCGCAAGAAT	CTTGCATTAC	ACTTAAAACT	TTAATATTTA	TTTTGCGCTT	109380
AGAAAGAATG	CTAACAGAAA	GGCCACTACC	TTGATCAATA	AAATAAGCAT	TCTTAAAAGC	109440
TCTAAATAAA	ACTCTATTAT	TAATTAAATC	AAGAACATTT	TTATTCTTAA	GATTCTTAAT	109500
AATTATTAAA	TCAAAACTTC	GATTTAGAAT	ATTATAAGCA	GAATCAAGTA	ACCCTTGAGA	109560
AAAATCATCA	TTTAAATTAA	AATTTGCAAC	CTTAAACTCC	AAAACACTAA	AATCTGTATC	109620
TACAACATTA	CAGGAAAAAA	ATAAAACACT	AAGCGGAAAT	AAACTCTTCA	AGTTGATACT	109680
TTGTCTCAAC	AAC TTCAAAT	ACAAGCCCAT	ACTTTTTTGC	AGCACTACTA	TCCAACCAAA	109740
AATCTCTATC	AGTATCCTTT	TCTATTTTAG	AAATTTTTTG	ACCCGTTTCT	TTTGAAATAA	109800
TATTATTAAG	TTCTTTTTTA	ACTTTATTTA	ACTCATTAGT	GTAAATCTCA	ATATCTGTAG	109860
CAACTCCCTT	AAATCCACTC	AAGGGCTGGT	GCAATAAATA	TCTGGCAAAG	GGCAGTGAAA	109920
ATCTATTTTC	TAATTTTGCA	GCCAAAAAAA	TTAAAGCAGC	AGCGCTAGCA	ACAAGCCCTA	109980
CTCCAAGTGT	AAAAACTTTA	GGCTTAACAA	AGCGAATCAT	ATTAAAAATA	GCAAATCCAG	110040
CATCAATGTC	GCCTCCTTCT	GAATCAATAT	ACACAAATAT	AGGCTTTTTA	AAATCTAGAG	110100
CCTCTAGCAA	TAATATTTTT	TCCTGAAAAA	GCCTGGAAAC	ATCCTTGGTA	ATCTCACCAG	110160
CAATAACTAT	TGATCTGCTC	TTTAAAACTA	ACTTCAATGA	TTTATCATGC	AAAACACAAG	110220
CATCATTATC	TTCTTTCCCG	GTCATAAAAC	ATCCCTTATA	CAAAAACATA	ATGATATATT	110280
ATAATTGAAA	ATAAAAGGTT	TTTAAATGAT	AAAAAAGCAC	AAAAATTAAA	CAATTGCACT	110340
TAATTTCTGA	AAAGCAAAAG	ACTAATAAAT	CTTTAATCAA	GCTTCATTAA	AGTTAAAAAA	110400
TACTCTAAAT	TTTACAAATT	AAGTAAAATT	AAAAAGGAGT	TTATAATGCA	CCATGAATTT	110460
GCGGTATATCG	GAGGGGGAAT	AGCGGGAAGC	ACCGTTGCTT	ACGAACTGCT	TAAAAGAAAT	110520
AAAAAAGTAA	TTCTTTTTGA	TAATGAAGAT	ACAAAAGCAA	CAATGGTAGC	GGGCGGGCTT	110580
ATTAATCCTA	TTATGGGTAG	AAAAATGAAC	ATTGCCTGGA	AAGAACCACA	TATTTTTGAA	110640
TTTGCAAAAA	ACTACTATCA	AGAAATTGAA	AAAACCATTA	AATCCAAATT	TTTTATAGAA	110700

AAAAATATCT	TTAGACCCTT	TACTACTGAA	AATCAAAAAA	ATGAACTGAT	TGATAAACTT	110760
GAAAAATAATA	AAAACATAAC	AAACTTTATT	TTAAAAATAC	AAGATGGAAA	AAC TTACAAT	110820
TTCTCAAACG	ACTCTAACGG	CGGAATGATA	ATAAAAGGCG	CCAGGGTTAA	TACAAAAACA	110880
TATATAAAAA	ATATTAAAAA	ATACTTAATC	GAAAAAAATT	CTTACATAAG	CAAAAATATA	110940
AACGAAAATA	AAATTAAACT	TGGAGAAAGT	TTTTTCAAAA	TAGAAGATT	TAAATTTGAA	111000
AAATTAATAT	TTGCAAAAGG	GTATAAGAA	AAACTCAAAG	GATTTTTTTC	TTATCTCCCA	111060
TTTGAGCCTG	CAAAAGGCGA	AATCATTATA	TTAGAATGCA	AAAAATTAAA	CTTTAAAGAG	111120
ATTTACAATA	GACACATATC	TTTAATTCAC	TTAAAAGGCA	ATAAATTTTA	CCTTGAGGC	111180
ACTTACGAAT	GGAACACTTG	GAATACACTT	ACAAATGAAT	GGGCAAAATT	AGAGCTATTG	111240
AAAAAATTTA	AAAAAATAAC	AAATCTAAAA	TGCAAGGTCA	TTGCTCAAAA	AGCACATATA	111300
AGGCCTTCAA	CTCTTGATAG	AGAACCTTTC	TTGGGAGAAC	ATCCTAAGCA	TAAAAATATC	111360
TTTATATTAA	ATGGTTTTGG	AACAAGGGGC	GTATCTATGG	CTCCATACTT	ATCTAATTTA	111420
TTAGTTAATA	ATATTGAAAA	AATTGACAAA	ATTCCAAATC	ATTACAATAT	TAAAAGATAT	111480
GCAAAATATT	ACAATATTTT	GGATCATTCT	TAAAATCAAA	ATTTTAAAT	CCATACATAC	111540
TGACAAACGA	CTACTATTAA	TATTTCTAAA	TTCATAAAAA	AATAATATAA	TGTTTAAGTT	111600
AAGCTAAAAT	AATTC TTATC	CAAAGAGAAA	CTAAGAGTGA	AACAAGATTT	AACAAAGCAA	111660
ATAAAATTAA	TTGACACTTA	CAAAACAAAC	CAGGAGAATA	ATCTTTGGGA	TTTAATATTA	111720
ATATCATAGG	AACTGGAGGA	ACAAGGCCAC	TCCACAATAG	ATATTTGTCA	TCCGTACTAA	111780
TCGAATACGA	TGGAGATAAC	TTTTTGTTTCG	ATTGTGGTGA	AGGAACCCAA	ATGTCTTTAA	111840
GGAAACAAAA	AATATCCTGG	CAAAAAATAA	AAATGATTTG	CATTACACAC	TTACATGCTG	111900
ACCACATCAC	GGGACTACTT	GGAATAGTAA	TGCTAATGTC	ACAAAGTGGA	GAAACAAGAA	111960
AAGAACCATT	AATAATCGCT	GGACCTGTTG	GAATAAAAAA	CTATACACAA	GCTAATATAA	112020
ATATGCTTAA	AATATATAAA	AACTATGAAA	TAATTTATAA	AGAAATAATC	ATAGATAAAA	112080
CCGAAAAAAT	AATATATGAA	GATAAAACAA	AAAAAATTGA	ATACACTAAA	CTAAACATT	112140
CAATAGAATG	TGTTGGATAT	TTATTTATAG	AAAAAGATAA	ACCCGGCAAA	TTCAACACAG	112200
AAAAAGCAGA	AGAGCTAAAT	ATTCTTAAAG	GGCCTATTAG	AAAAGCCCTA	CAAGATGGAA	112260
AAGAAATATT	GGTAAACGGA	AAAATTATAA	AGCCATCAGA	AATACTTGGA	AAATCTAAAA	112320
AAGGACTAAA	AGTTGCATAC	ATTACAGATA	CTGGTTATTT	TAAAGAACTC	ATACAGCAAA	112380
TCAAAAATTT	TAACCTTGTA	ATAATTGAGA	GCACATTTAA	AAATGAGCTA	AAAAAGAAG	112440

CCGATAAAAA	ACTTCACTTA	ACAGCTGGCG	GGGCTGCAAA	TATTGTCAAG	CAAGCAAAAG	112500
TTTTACAAAC	AGGACTTATC	CATTTTAGTG	AAAGATATAC	ATTAAGAAAA	GATCTTGAAA	112560
ACTTACTAAA	GGAGGCAAAA	TTGGAACATC	CAGACGGAGA	AATTTTTTTTA	ACAAGAGATG	112620
GAATGAGGCT	TGAAGCAAAAC	AAAAATAACT	TTATTATTAA	ATAGGAGGGT	ATATGATAAA	112680
TGTAGAAAAA	GTTACTAAAA	TGTATGGGCC	ATTTACAGCA	CTATTTAATG	TTAGCTTTAA	112740
GGTTGAAGAA	GGCGAAGTAC	TTGGTATACT	TGGCCCAAAC	GGAGCCGGAA	AGTCCACATT	112800
AATCAAAATC	TTAACATCAT	TTCATTATCC	AAGCAAAGGT	AATGTAAAAA	TTTTTGGAAG	112860
AGACATTGTA	GAGCATTCTG	AAGAAATACT	ACAGCAAATA	GGATATGTTC	CTGAAAAACT	112920
AGCTCTTTAT	CCAGAGCTTT	CTGTTAAAGA	ATATTTAAAG	TTTATATCAG	AAATAAAAGG	112980
TGTTAAAAAA	TTAAAAAAG	AAATTGACAG	AGTAATAAGC	ATATTCAAAT	TAAAAGAGGT	113040
TGAAGATAAG	CTGATTTCTC	AACTTTCAAA	AGGATTTAGA	CAAAGAGTAG	GAATAGCTGG	113100
CGCTTTAATA	AACAATCCTA	AACTTGTAAT	ACTTGATGAG	CCAACAAACG	GTCCTTGATCC	113160
AAATCAAATA	ATTGAATTTA	AAGAATTTTT	AAGAGAACTT	GCAAAAGAAA	GTACAATATT	113220
ATTCTCTTCG	CACATACTAA	GCGAAGTAGA	ATCTATTTGT	AAAAGAATAA	TTATTGTCAA	113280
CAACGGAGTA	ATTGTTGCTG	ATGACACAAA	AGAAAATATT	ATTAAAAATA	AACTTAAAGA	113340
GATTGAAATA	GAATTAATAG	TTTCAAAAAA	ATCTGAAAAT	GAGAAAAAAA	TTTTCAACAG	113400
CAAAAATGAT	ATTTTTTTCAT	TAATAAAGCT	TGAAGAACAC	GAAAAAGACT	TAAATATTTT	113460
ATTAAAACTA	TCTCAAGGCA	AAACAGAAGA	AGATCTCTTT	AGCTACATAG	TAAAAAATAA	113520
TATAATCTTA	AAAGCAATGA	TTCCAAAACA	TGAAAGCCTT	GAAAAGATAT	TTAGCAAATT	113580
AACCAAGGAG	AGAGAAAAAT	GAAAATAGAT	TTAAAGCAAT	CTTTATCGCT	TTCTAAAAAA	113640
GAACTAAAAA	TATTATTTGG	AACCCCAACT	GCATACGTTG	TGATGCTATT	TTTTTTAATA	113700
TTCATAAACT	TTTCATTTAT	TTTTTTATCA	GGATTTTTTA	TTAAAGACAA	TGCATCTCTT	113760
ACCTCTTATT	TCTCTTCAAT	GCCTATTATT	TTAATGTTGG	TACTGCCAGC	ACTTAGCATG	113820
GGAGTATTCT	CAGAAGAACA	CAAAACAGGA	AGCATTGAAC	TTCTTTATGC	TCTACCGCTA	113880
AGTCCTCAAG	AGATAGTCTT	GGGCAAATTT	ATTACGCTTA	AAATATTTAC	CTTAATACTA	113940
TTCTCACTTA	CCCTACCTCT	TACAATAATG	ACAATTTTCA	TGGGCGAATT	TGATCTTGGG	114000
ATAATATTGC	TTCAATATCT	AGGAATAATT	CTTTATTCTC	TTTCTGTGCT	AAGCATGGGA	114060
ACATTTATAT	CCTCCATTAC	AAAAAGCCAA	ATAGTCTCTT	ACATTCTTAC	CGTATTTACA	114120
CTGATATTAA	TACTATTTTC	TGGGAAATTG	GTTATGATCT	TTGGAAAAGA	AAATATAATA	114180
GGAGAAATAC	TTAATTTTGT	TTCAATAACC	AATCACTTTA	GCTATTTTAA	TATGGGTATA	114240

TTAAACTTAT	CAGACTTTAT	TTATTTTATT	ACATTCACAG	TCACATTTCT	AATACTAAGC	114300
ACACACAGCA	TAACACTAAA	AAAATGGAGA	TAAATTTATG	AAAAACAAAG	AAAATGAAGT	114360
TTTAAACCTA	ACTTTGAACC	TTACAATAAT	CTTTTTGATT	TTTGTGAATA	TATCTATTTy	114420
CATTTTTTAA	ATAGACTTTA	CAAAACACAA	AGCTTTTACA	ATATCTAAAG	TTACAAAAAA	114480
TTTGTTCCTA	AGTGCAAATG	AAACAATATA	TATAACATAT	TACAATTCAG	GAAGCCTTGA	114540
AAACTATTTT	GCTTTTCCAA	ACCAAATAAA	AAATTTTTTA	ATAAGTTTTT	CTGATGCTTC	114600
AAAATGTAAG	GTAATTTATA	AAGAAATTGA	CGCTGATAAA	ATTTC AACAC	CATTAGAGCA	114660
CATTGGTATT	CCCTCTCAGC	AAATCGACTT	AAGAGATATT	AATCAGCTTT	CAATACTCAA	114720
AATATACTCA	GGAATTGAGA	TTATTTACGA	GGGAAAAAGA	GAAGTAATAC	CGGTGTGAAC	114780
AGAAATCAGC	AATCTAGAAT	ATGACCTTGC	AAATGGACTT	GACAAACTAA	TAAATAATAC	114840
CAAAAAAGTT	TTAGGACTTG	CTTTTGGAGA	CAGCACTTTA	AAAGAAGCAC	ATAAAAACTT	114900
TTCCGAAATÀ	ATGAAAAAAG	CATTTGGAAT	TGAAATAAAA	GAAATAGATT	TAAAAACTGA	114960
AAAATTAGAA	GACATTAGAA	AAGATATAAA	TGGATTATTC	ATTATTGGCG	CTAAAGAAAT	115020
TGACGAAGAA	ATTGCAAAAA	AAATTGACGA	TTTTATTGTT	AATGATGGAA	AAATATTTGT	115080
TGCAACAAGC	ACAATTGACT	ACAATCCTCA	AAATCCATAT	GGCATAACTC	CTATTAAATC	115140
CAGCCTATTT	GATCTATTTG	AAAGTTATGG	GATAAAATAC	AACGATAATA	TTATTCCTGA	115200
TAAAAGAGCG	CCCACAATCT	TTTTGGGTGG	CAATTTCCAA	ACTTACTATC	CATGGATCTT	115260
AATAGACAAA	AGCAATATTG	TAAAAAAAGA	CATGCCATTG	CTTAAAAATT	TTTATACCGC	115320
TACAATTCCT	TGGAGCAGCT	CATTAGAACT	TATAAAAAAA	GATGAAACAG	AAGTAAAATT	115380
TTTACCTCTA	TTTGCAAGTT	CCAAACAATC	ATGGCAAGTT	AAAGAACCTA	ACCTTTCAAA	115440
CATATCTTTG	AATGCATTTG	AAGTTCCAAA	TAAATTTGAA	GAGAATAAAA	CTAAAATACT	115500
AGGATATGCA	ATTGAAGGAA	AAATTAAAAG	TCCTTATAAA	GATCAATATT	CCAAAAATTC	115560
TAAAAATAATC	CTAACAGGAT	CAAGCATGAT	ATTTAGCGAT	TATATGTACA	ACGGGTCTCC	115620
ATCAAACCTT	GAACATCAG	GAAGAATTTT	GGATTATTTA	ATGCAAAAGG	AAGAATTTTT	115680
TAATATTAAG	TCCAGAGAGG	TACGAGCTAA	ATTAAAATTT	GCAAGCTCTT	CAAACGAAAT	115740
GGTCAATGCA	AAGTTTTTCAT	TAATAATTGT	TAACCTAATT	ATTCTTCCAA	CAATAATATT	115800
AATATTTGGA	CTTGTTAGAT	TTACTAGAAA	AAGAAAAGCA	AATTAATAAG	AACAAAGGAG	115860
TGTTTATGAC	AAAACCAAAA	ATAFTCTCAA	TCAATAAAGA	AAAAATAAAA	ATATTGATAA	115920
TAGTAGTGTT	AACATCTACA	TTCTTATTGG	GAATAATTTT	TTCAAATGAA	AATAAAGTAG	115980

CAAGGATCCT	TGAAGAAAAA	TTTTTTGATT	TTGACTTTAA	TTTAATTTCT	AAAATTGAAA	116040
CAGAGCTTGA	AGGAACGCTA	ACAAAACCTG	GCAAAGATTG	GATTTTAAAC	TACAATAAAC	116100
AAAATATTCC	TGTTGATAAC	AAAAAAGTCA	ACTCTCTAAT	CAAAGCATTA	GACGAGCTTC	116160
AAAAAAACAA	GCTTGTAAGT	AGAGATCAAA	AAAAACACAA	GGAACTAGGA	ATTGGAGAAA	116220
ATCCAAGCTT	TAAATTATTT	GACAATAATA	ATAAGCTGTT	AACAGAAATT	TTTGTGGAA	116280
AATCAGGAGA	AGGCGATTCA	AGACTGGCAT	ACATTAAAGG	TAGTGACGAA	AATGTTTACT	116340
TAACAAAAAA	CATTTTCTTA	TCATACAAAG	GAAATTCTTA	CAATACATTT	TCAGATACTA	116400
CATTGTTCCA	AGAAAAAAAC	ACAAAATTAG	AAAATTTATC	ATTCAAAATA	ATAAGAAAAT	116460
TAAACAAGGA	AAATGAAAAT	AACATAAATA	ATAACTATGA	GATTATCAGT	AAAGATGGCC	116520
TTTATTTTTT	AAATAACCAA	AAAATGACAA	AAGAAAGGCC	TTTAAATATT	ATTGCTGAAT	116580
TTAAAGCTGA	CGGACTTGAA	ATTGATAAAT	CTAAATAGA	TGATTATAAT	CTTCAATACA	116640
AAATTGAAGT	CAAATGGAGC	AATAAAAGTG	TCAATAATAT	TGAAGTTTAT	TTTAATAAAA	116700
ACGAAGAAAA	TGACAAAGAC	ATATTAATCA	AAAAAGATAA	AGATGAATAT	TACTACACGA	116760
CTAGCAAATG	GACTTTTTTT	GATGTATTCG	ACTTAGAAAA	AAAATTAACA	GAAAAAGATG	116820
ATATTTCTAG	CAACGATAAT	CAAGAAGATC	ATCATGAACA	TCACAACAAT	GCAGATTAAT	116880
CTTGCTATAT	ATAAAAAGCA	TTAAAAGAAA	AACATATAAA	AATAAATATA	ATTAAAATAT	116940
ACCATGACAA	AGACAACATT	TTATCAAAAG	ATAAGATGTT	GTTTGGCTTT	ACTGATATAT	117000
CTTTTAAATT	AAAATTAATA	TCAAACAAAC	CGCTCAGTTA	TCAAATATTA	ATTTTAAGAA	117060
TTTTTTATAA	AAATAGAACT	TAAACGAATG	GATTTTCAAC	CTTTAGTAAG	TAAAAATTTA	117120
ACTTTTTTTA	AAACTTCATA	CTCTTGTTTA	ATTTTAAAAA	TATTTCTATT	AGGATTTAAC	117180
TCAAGTTCAC	TTTCTACCCT	ATCAATAAAA	TTTATTAATA	TGGCTTTTTT	ATCGCTATTT	117240
AGCACAACAT	TACTTAAAGA	TTCTTCAGCA	TTAAATTTGC	TTACAACCTT	TTCAACATCA	117300
TTAAAATTAA	AATTAGAAGG	AAAAGGGTCC	TTAGTTATGC	TTTGAGAGCT	GGGCTTTTCA	117360
TAATTAGCAA	CATCATTAAT	TTGGTCTTGA	TCTGATTTAT	CTCTCAAATT	ATCATGCAAA	117420
TTTTTCTCTC	CAGCATTAGA	AAAAGAGCCT	TCGGCAACCG	TAGAATCCAG	ATCCTCTTTT	117480
AAAATATTTT	CAAAATCATC	AAATTCTTTT	GACTCAAAAC	GTTCAACAAA	GCTTTTAAGA	117540
ATCGCACAAAT	TATCACCATT	CTCATGCTCA	ATGTTTTTCAT	TCTCATTAAT	TAAATCCAAA	117600
CGTTGTTTAT	TATTATCTAC	AAAGCTCTCC	AATTTTCGAG	AGTTATCGCA	ATTAGCAATA	117660
GAATAGGGAT	CTTCTGCTCC	CACTTCATGT	TCCAAAGAAG	AACTATTATC	CAAAATTTTA	117720
TTAACAGAAT	CTAACAGTTT	ATCTGGCTCT	GTAGTAAAAC	TTTCCAATTT	TTTACTGTGA	117780

TCGTTATTAT	TATTTAACTC	CATATTTCCCT	ATGGACACAT	TATCACTGTC	AACATTAATA	117840
TCCTCTTTTT	TAAGAATATT	ATTAGAAAAA	TTATCAGCCA	CATGTAAATT	GTTAGTCAAA	117900
TTCAAATCTT	TATCTCTATT	AAATTTAACA	TCGTGCTGAA	CTTCTTCATT	TTCTTTTGGG	117960
GTATAATTGA	TACCTTCAAG	CAGCGCATCT	AATTCTTCTT	GACCAATAGA	AATATTAGGC	118020
AAATTATCTT	CTTTTTTTGA	AAATGAATCA	TCCGTCTCGG	ATGATTTTTT	TTCAAAATCC	118080
TTCGAATCAT	AAGAAATAAA	ATCCCTTTGA	ACAAACTTAA	GACTATTATC	AAGCATTTCT	118140
TTTTCAACTC	TCTCAACACG	AAGCTCAACA	CCCTTTAAAC	AATTAATCAG	CTCGCCGTGC	118200
CTTCTTTTTA	AGATGTTATC	GAGTTTTAAC	AGCTTTTCAT	CCAAAAAAGT	TTTAAGATTA	118260
TCTGGGGCTA	TAGAAACAAT	ATCAACAGAC	TCACTTCTTA	AAAATACCTT	TTCAATATCA	118320
AAATTAACCT	CCTTGGGACC	TTCTTTGATA	AAAAAAACAC	TTTCCTTCAT	GCCAAACTTG	118380
CTCTCCAAAA	CTACTCAATA	AAACTACAAT	CTAAAGCAAT	TTTAACTACT	TGTAAATATA	118440
GTATATTAAG	AATATAATTA	CAAGCTATAT	GACTATTTAT	AAAAAAATTG	CAATGTCTTT	118500
TTACTCAGGA	ATACTAAGCT	ACTTTATAAT	AGCTCCCAT	TTTGAGAGA	GAGGATTTGT	118560
TAATTATCAA	AAATTGGATA	ACAACTTAAC	ATTAATAAAA	AATCACATCG	AAAAACTAAA	118620
AGAAATTCAA	AAAGAATTAA	AAGCAAGATA	TATTAACCTA	CAAGTATCTA	AATCGGAAAT	118680
TCTAAAAGAA	GCTAAAAAAT	TGGGCTACTA	CCCAAAAAAC	TCAACAGTAA	TAAAAACCAA	118740
CAATAATAAA	GATCAATATA	ACCAAGGGCA	AATATTAACC	TTACAAAAAC	CCCTTTCCAA	118800
GAATCAAAAT	TTTTACCTTA	TATCAATAGC	AATAGGTTTA	ATTTATTATT	TTTTATCAAG	118860
CTGCATTATC	CAAACCAAGA	AAATTACAAA	AATCAATAAA	CTTGCTTCCA	ACAACTCTAA	118920
GGATTAGTCT	TTATTGAAAA	TATTTATTTT	TAAAAATACA	ATATATTTAT	TAATTAATTT	118980
AATTTGTGCA	TCATTTTTTT	GCGTATCGTT	AGTAAATCTT	TTTTCAAATG	AACAACAGTA	119040
TACTCCTTTT	GTTAAACAA	ATGTCATAAA	AAATTACTTA	CAATACATTG	GAGTATATAA	119100
AAGTATAGAA	AGATATGCCC	TGATACATGA	CTTTAACCCCT	AAATCAAAAT	TAGAAAAAGA	119160
TTGCTTTTTG	AAGCATATAG	CTGGCAATTC	ATATATAATA	TACAAAACAA	AAAATGAAGG	119220
AATGCTGTGG	GGCGATCATC	GATACTCTCT	GCTGAGCAAA	GGAAAGCCAA	CTACTAAAAT	119280
AATTTTTCAA	AAAATATTTA	ATACTTTAAA	AATCTCAATT	CCAGGCGCCC	TACTCTCTTA	119340
TATTGCGGCA	ATAATCCTTA	TTATAATTTG	GAAAATTTAC	ATAAAAAATA	ATCTAATAAA	119400
TAATATTCTA	GAATATTTAA	TGCTATTGCT	CCACTCCATG	CCAAGAAACT	TAACAGTATT	119460
TTTAATACTG	TCTTTAATAT	ATTACCTTAA	TTTAAATCCA	AAAAATTTAA	TAATGGGTGG	119520

ATTTGCATGG	TTTTTTTCAT	CTTCATATT	TAATTCTGTA	ATTTTTAAAC	ATCTCTTGA	119580
CAAAACTTTA	TCAGAATTTT	ACATAAAAGC	TGCAAAATCA	AGAGGAATAA	ATAAATTGCA	119640
AATAATCTTA	AAACATGCAT	TAATTCCATC	AATAACACCA	TTACTCACAA	ACATGAGACC	119700
TATTATTACA	ACAGCTTTTT	TTGGAGCATC	AATGATTGAA	TCAATGTTTG	AAATTGATGG	119760
AATTGGGGCC	TTATATTTAA	ATGCTTTGAA	ATTTAACGAT	TATGCTATTT	CTAAAGATTT	119820
GATTTTTATT	GGCGTTTTCA	TTATGCTTAT	TCCAAATATA	ATAACAGATA	TACTAATTTA	119880
CAAAATTAAC	CCATATAAGG	ACACTCTAAA	CTAATGAAAA	CAGATACAAT	AATAAAAAAA	119940
ATTTATATCG	TACTCTTTAA	TATATTTATT	GTGTTGCTAA	TTATTACTCC	GTCATTGGTT	120000
AATGAAAATT	CAAAAATTGC	AATCTATAAA	AAAGATCCAA	ATAAAGTCTA	TTTAAATCTT	120060
ATTAAAAATG	TACCTATGCC	ACCCACAAAA	GACAACCCAT	TAGGAATCGA	CAAAATGGGA	120120
AGAGATATTA	TGGCAAGATT	AATAATTGCA	ACCAGAAACT	CTATTTTACT	TTCACTAAGC	120180
TACGCAACAA	TTTCTGCAAT	AATTGGAATC	TTTATTGGAA	CAATCATTGG	CATGTTTAGT	120240
TTTGAAATTT	GCATGCTGAT	TTCAAAACCA	ATTGAAACAT	TGCAAACATT	ACCTTTTTTT	120300
TACGTTGTGT	CTTTAGTTTT	TTATTACTTT	TTAAAACAAA	AAACTTACAA	TATGCTTCAA	120360
ACAGCAACAC	TATTAGCATT	GATTCATGGA	TGGATTAGAT	TTGCTTTTAT	TGCAAGAAAC	120420
AATACATTAA	TAATAAAAAA	TTTAGATTAT	ATTAAAGCCA	GCGAAGCTAT	GGGAGCAAGC	120480
AAAATTAGAA	TAATATTGTA	TCATATTTTT	CCAGAAGTAT	TCTCATCAAT	ATCATCTATA	120540
ATCCCATTAC	AAATGGGAAG	AAGTCTTACT	ACTTTTGAAG	TAGTAAGTTT	TTTACAAAAA	120600
CAAGATAAAA	ATCTATATCC	CAGTCTTGGA	GAAGTCTCA	ACTATATGCA	AATGGGCAAT	120660
AAATATCTAT	GGATATGGAT	CAATCCCTTA	CTCATATTAA	TAGGCATAAA	CATAATACTA	120720
GCAATTATAA	ATTTTAAGCT	AAGAAAAAAA	ATGAAACATT	TAATATCATC	TTAAATAAAA	120780
AATTAACAAA	CTCTTGGAGC	AAATTTTTCT	AAAAACAAT	TATCACAATT	TACATTTCTA	120840
GAAGTACAAA	TTTCTCTTGC	ATGCTTATTA	ATAGCCATAG	AAAATCTATA	CTGCTTACAA	120900
GGCTTTATTC	TTCTTTTTAG	ATCCAATTCA	ATCTTAATAG	GAGAACTTTC	CAAAGAAAGA	120960
GCATGTCTTG	TAATAACTCT	ACTAAAATGA	GTATCTACAA	TAATTGCGGG	TTTATTGTAA	121020
ACAGATCCAA	GAATAACATT	TGCCGTTTTT	CGACCTACTC	CAGGTAGCTT	AATAAGATCA	121080
AAAATATTAT	TTGGAATAAC	ACCATTAAAT	TTTTCTAAAA	TATCAATAGA	GCAATTCACA	121140
ATATTTTTAG	CCTTTCTTGA	ATAAAAACCA	GTCTTATAAA	TTAATTTTTT	AACATCTCTC	121200
ACATTTGCTC	TTGATAAACT	TTCAAAATTC	TCGTACCTTT	CAAAAAGGTA	TGGAGAAATT	121260
TTATTCACCA	AATTATCTGT	TGTTCTTGCA	CTTAAATATA	CCATTATTAA	AAGTTCATAA	121320

TTGTTTTTAT	AATTTAAAA	AGGTTTAAACA	TCAGGATATC	TAAATAAAGT	TTCATCAACA	121380
ATCAAAATCAA	GATTAATCAT	AAAAAAATTA	TAAAACATTA	TAAACACAAA	ACAAAAATAA	121440
AAAATATACA	AAGTAAAGGT	ATCTAGACTT	TATTGACAAG	GATTTTTTCAA	AATGATATAC	121500
TCATCATTAG	AATTTTAAAT	GCACCAATAG	CTCAATTGGA	TAGAGCAACA	GACTTCTAAT	121560
CTGTAGGTTT	TAGGTTTCGAG	TCCTAATTGG	TGCGCTTCAT	TCGGGATGTG	GCCTAGTGGC	121620
TAAGGCACCT	GCTTTGGGAG	CAGGGGATCG	TGAGTTCGAA	TCCCACCATC	CCGAAAAAAT	121680
ATTAAAAAAG	CTAAAAACTT	TTGTTTTTAG	CTTTTTTGGT	TTTTTAAACGA	TTTATACAAA	121740
TTAAATCTAA	CTGTAAAGTT	ACTTAACTTT	CTTTAAAGTA	TTTACATCTA	AAATAACTAG	121800
ACTTTTAAAC	TCATCTTGCA	AATAAATAAA	ATTTTTTCTC	ACAGAAAAGC	TAGTAAAAGG	121860
CATAATTTTA	TTCTCTGAAA	GAATAAACTC	ATCTAAATTT	TTAGGAGAAA	ATTTGGCCAA	121920
TCTCCAATCA	TTACTACTAT	CTTTATCCCT	AACAGCTACT	AAAATCATTT	TAGAATCAAC	121980
ATAAAGAGAT	GAATTTTTAT	TAATCTCAAA	ATTAGACTCT	GATACCACTT	TTAAATTTTC	122040
AAGTTTATCA	AGTATCTGAA	GCTTAGCTTT	TCCTGAATCC	ATTTTAATAA	CAACCAAATC	122100
TTTTTTCACGT	TCATAAATTC	CATACCGCTG	AATGCCTTGC	TGAGTGCTTT	CTTTAAGCCT	122160
AACACCAGTA	TTTAAATCAA	TAAGTTGAAG	AGTTCCTAAA	TTTGTAAATG	GATCAATAAC	122220
CTCTAAAAAT	ACAGGACTAC	TGGAATCTAT	AGACATAGTA	GTCAAATCTT	CATTCAAAGA	122280
AGTAACTTGG	TCTTTAACCT	GAGGCTTAGT	CTTTTGCAAA	TTAACATCTT	TATTAAGTGT	122340
CTCCTCTTTT	GAATCAATGT	CTTTATAAGA	AGATTTATCT	AACGGTGATA	ATTCTCCAAC	122400
ATTGTTATTA	GACTTGAAAA	TCTTATCTAA	TTTCTCAACC	TCAGAAACAG	GTTTAAATTC	122460
TTTTTTGCTA	TCTAATTTTT	TAACCTCAGG	TAATTTTTGA	TCTTCTGGCA	TCATAAGATT	122520
TTCATCATTA	TTCAAATCGC	CTAAGCTTTT	CTGTGACTTA	CCCTTGGTTA	TTTCTTCTTC	122580
CTTGGCTTTA	CTTTTTTCTT	TGCTAGAAGC	TTTAGAATTT	AATTCTCGAT	CAAGATCCAA	122640
GGCTTTACCA	TCTTTACTTG	CTTTATCATC	TTTACTTTTT	AAAAGCTTTT	CATCACTTTT	122700
TTTGATTTCA	ATTTGCTTTT	CAATTTCTCT	TTTCTGATTT	TCATCACCAG	TTTCTTTAAG	122760
CTGCTCCTGC	AAATCTTCCA	GGCTCTCTTT	TATTTGTAGT	TGCTTATCAA	CTTTAGGAGA	122820
ACTTACATCA	CCAGGCTTTG	GTAAATTCTT	TTCCTTGTTA	ATTTTCGTTAA	TATCCTCTTG	122880
AATTTTCTCT	CTAACAGTAT	TTCTTTGAAC	ATCTAAATTA	TCTTCAGCAG	AATCTAATTT	122940
TTGCTGAGCT	TTATCAAGAT	TTATTGCCTT	TTTATCTAGC	TCTTCCTTTT	GTTTCTTTTT	123000
AGCATCAACC	TGACTTTCAA	TCTCTTTTTT	ATGCTCTTCA	TCTGTAGCTT	TTTCAAGCTG	123060

ATCCCTTAAA	TTTTCAATAG	TTTCTGTAT	ATTGGAATCA	CTTTCATGAA	TTTGTCTAA	123120
TTCAATATCA	ATTTTATCTT	GATCTGCCTT	ATGAGTTTCG	CCTTGAATAT	CTGTAATATC	123180
TCTTGCAAAG	TTAACACCTG	CTTCATTTTC	ACTTAAAAGA	GCTGCCACCA	CCTTATCTGT	123240
AACTAAACTG	TCAATATCAA	TGTCAGACTC	AATATTTCCA	GACAAAATAT	CCTTTTFAAG	123300
AGGAATAAAT	ATTTGTGTCT	TTCCAGCCCC	CTGACTATAA	ACCCTAGAAA	GACCTGCATT	123360
TTCTTTACTT	AAAGACTTTA	AAGCAGCCTC	AATATAAAAC	CCTTTATAAT	AATCCAAATC	123420
TCCTCTATAA	ACAGCATTAT	ATATTGTAAT	AACCTTAGCA	ATTAATTCTG	CACTAGACCT	123480
GTCATAATCG	AAAGACTTTA	TTAAATACCC	TGTAAGAATT	CTTCTTAAAT	TCAATATACT	123540
GTCAAGCTCT	GACTTACTAC	CAATAGAAAA	AACATCAACG	CTTGCTTTTT	TATCTTGATC	123600
ATCAATAAAT	CTATTAATAA	AATATTTACC	ATAATAACTT	GAGTTGCTAT	TGGAATTGGT	123660
CAACGGTCTT	GCTAAAAACT	CCCCAATACC	CACATTTTGT	TCATATGTAT	TTGTAGAATC	123720
ATAAGGGCCT	TTATAATTTA	CAAACCTCAAG	ATCCATATTA	ACAAAGTCCT	TTAATTTTTC	123780
CCTATCAACT	TCTCTTGCAC	TAACAGGAAA	TCCATTCAAG	AAAATAAGAA	AAAAACTAAA	123840
GATTAGTAAC	ATTTTTTTCA	TAAAAGAAAT	TCTCCTATAA	ATTTAATTAT	AATCTACCTT	123900
ACCAACTAAA	TTCACCAATT	TAAACCTAAT	TTTAACACAA	TCGGATTTAT	ATGCAAAACA	123960
GCTAATTCAA	TCTTGGGGAC	TTGAAATATC	TTGAATGCTT	GAAATAATTT	CATTAATGTG	124020
CTTGTTAATA	TTTTTCAGAA	AATTTTGACG	CCTCATAAAC	ATTAATAACT	TTATATACTT	124080
ATGCTTACAA	AGTTTAAGAA	GTTTCGGCTGT	TGAATTTAAA	ATTATATTTT	TCAAATTTAA	124140
AATCACAGAA	TAAGCTATTT	CCAAATTACC	CTGATAATTT	AACACTATAT	TTTGGATGTA	124200
ATTGTATATA	ATACAATCTT	GCCCATAATA	ACTTTCAAAT	TTAGTAATAG	TTTTGACTAT	124260
GGTAATTATA	TTATTAATAT	CTCTCTCTCT	AATTGCCAAA	TTAATAAAAC	TCTTAATTGC	124320
AGCTTTATCA	TTAACATATA	CAAAATATCT	ATAAGCTTG	CCTCTTGAAG	CAATTCTATC	124380
AAACCTCTTT	TCTGACAAAA	GCCACACATT	AAAAGCTTTT	TCATATAAAT	TTAAATATTT	124440
TGCGGGAAAA	TTTCCATGTT	CAGCTGCAAG	AAAATCTAAA	ATCTCTAAAA	ATTTATTATA	124500
AGCAGCATCA	CTATAATTAG	AATTAAGGAT	CTCTTCAAGA	AGTAAATAAT	TTTTATCATA	124560
ATCTACAAAA	CGCTTTTCTT	TTAAATCTAT	TCTGCCTTGA	TTATATTTAT	TACCTCCAAA	124620
ATGGTTCCAA	ACGCTAACAT	CAAGCTCTCT	GTCGTATTCG	TAATAATAAG	TACTAAGAAC	124680
CCAATCATCT	CCCCCTAAGA	CAATAATACC	TGAATTAGCA	CTAAGTCCTG	ACAAGCCCTT	124740
AAGTTTAGCA	TCTGAAAAAG	ATATATCAAA	ATCCTCTTCA	AAAGACAAAA	TTCCATTTGA	124800
TTTGTTGAA	ATTAAAAGAT	TTCTATTATT	ATAAAGAACA	GCTTCTCAA	TCTCAAAGTT	124860

CATTTTTTTT	TTAAATTTAA	GCTTAAAGTT	TCTATTCAAA	TCAAAAAACA	TACCAACATT	124920
TGAAATTGCG	ACATATTCAC	CATTAAACTT	TTCAAACAAA	AAATTAATAG	GAAAATCCAA	124980
CTTAATAGAG	CTAATAACTT	CCCCAAAATC	TCTTCCATAA	GCAACAATAT	GCCCTGACTT	125040
ATGTCCAACA	ACTACCTCTT	TTTTTGTATT	TATCATTAAC	AAAAAAGGGA	AAGCAACTAA	125100
TCTGAAAAAC	CATTTCTTAT	TACCCAAAGA	ATCAAACAAA	AATATTTTCAT	CATTTTCACT	125160
TGCAATACAA	AAATCCCCAT	TATCAAAAAC	TACAGGAGAA	GTAGCAGGCC	TACCACCTAT	125220
GTCAACCTCA	AACATTTTTT	TACCGCTATT	TAAATCAATA	GAAACAACCT	TTTCATTAGC	125280
AAGAGGAATT	AGGATGTTAA	CATTTCTTAT	TGCAGGAGAG	CTCAAAGGTG	AAAAATCAAG	125340
CTTATACTTC	CAAACGAGTT	TTCCTCTTCT	GATCTTTTGA	ACTTCATTTT	TAAGTGTAA	125400
AACATAATAC	CCATTATCAA	AATCTTTTCAA	AAGAAAAGGA	TATGGCATT	TATTTAATCT	125460
ATAAGAATAT	TTCTTCTCAA	ATGACATTGT	ATAAGTAGTC	AACCATCTAT	CTTTTGTTAA	125520
AACTGTAATA	GTGTCACGTT	TTTCATCAAT	AATTGGATTG	CCTGCAACTT	TGCCAGTTAA	125580
TGCTTTTTGA	AAATATAAAT	TAATATCAGA	ATAAAGCCTT	AAAAAAGAAG	CTGAAAACAC	125640
AAATATGAAA	AGTAGACCTC	TCAAAATAAA	AAACCTTTTG	AGTTTCTAAA	AACTGCTAC	125700
TAAAGCCTAA	AACCAGCATT	ATTGCCATAA	AGATTATTTT	TCAGATCCCT	AATCTTAGCA	125760
TCATCAACAT	ACTCAGAAAA	AGTCATATAT	CGATCAATTA	TTCCGTTAGG	AGTAAACTCT	125820
ATAATCCTAT	TAGCAACAGT	ATCTATAAAT	TGATGATCAT	GTGATGTAAA	AAGAACAAC	125880
CCTTTAAACT	CTTTAAGCCC	GGAATTTTAA	GATGTAATTG	CCTCAAGATC	TAAGTGATTT	125940
GTGGGTGGT	CCAGTATTAA	AACATTAGCT	CCGCTAAGCA	TAGCCTTAGC	AAGCATGCAT	126000
CTTACTTTTT	CTCCCCCTGA	GAGAACATTT	ACCTTTTTTA	AAGCTTCATC	TTGGCTGAAA	126060
AGCATTCGAC	CTAAAAATCC	TCTAATATAA	GTTTCATCTT	GTTCTTTTGA	ATACTGACGT	126120
AACCAATCGA	CTAAATTTAA	ATCTAAATCA	AAATATTTT	CATTATCTTT	ATTAAAATAC	126180
GAAAAATTAA	CGGTAGATCC	CCATTTCATA	TGACCTTTAT	AATTTCTATC	TTTATTGTA	126240
ATAATATCAA	ACAAAAAGT	TGCAAACATG	GGATTTCCCA	AAAAAACAAT	CTTTTGCTGA	126300
GGTTCAACAA	TAATACTAAA	TTTATTTTAA	ATTAAATTCC	CTTCAAATTC	TTTTATTAAA	126360
TTTTTAATTG	TAAGAACATT	CTTGCCAAGT	TCTCTTTCGC	TTTTGAAATT	AACATAAGGG	126420
AACTTCCTTG	AAGAAGGCTT	TAAATCTTCA	ACCTTTATTT	TTTCAATCAA	CTTTTTCCTT	126480
GATGTTGCTT	GCTTAGACTT	AGATGCATTA	CTAGAAAATC	TTTGAATAAA	TGTCTTAAGT	126540
TCAGCAATTT	TATCTTCAGA	TCGCTTTTTA	GCATCTTTTA	GTTGCTTGTT	TAAAATCTGA	126600

CTTGTTTCAT	ACCAAAAATC	TTAATTTCCA	AGATACACTT	GAATCTTGCC	ATAATCAATG	126660
TCAACAATAT	GAGTACAAAC	TTGATTTAAA	AAATGTCTAT	CGTGAGATAC	AACAATAACT	126720
GTATTTTCAA	AATTAATTAA	AAACTCTTCT	AACCATTTAA	TAGATTGTAG	ATCAAGGTTA	126780
TTAGTAGGCT	CATCAAGAAG	TAATACATCG	GGATCACCAA	AAAGTGCTTG	AGCCAAAAGA	126840
ACCCTAACTT	TTAAAGCCCC	TTCAACATCA	CCCATTAAAT	TATTATGAAT	TGCCTCATCT	126900
ATTCCAAGAC	CTTTAAGAAG	AACCGCTGCA	TCAGATTGAG	CCTCGTATCC	TCCAAGCTCT	126960
GAAAATTCTG	CTTCAAGCTC	TCCAGCTCTA	ATTCCATCCT	CATCAGTAAA	ATCAAGCTTA	127020
CTATAAATTT	CATCTTTTTT	TTTTTGAACA	GAATAAAGTC	TTTTGTGACC	CATAATAACA	127080
GTATCGATAA	CCTTATATCC	ATCATAAGCA	AATTGATCTT	GTTCAAGAGC	TGCTACTCTT	127140
TGATTTTTTG	GGATAGATAT	TTCACCCTTA	CTAGCTTCAA	TCATTCCCCC	TAATACTTTT	127200
AAAAAAGTGC	TTTTTCCTGC	CCCATTAGCA	CCAATTATTC	CATAGCAATT	TCCAGGAGAA	127260
AATTTAATAT	TTACATCTTT	GAATAAAACT	CTCTCTCCAA	ATGCAACTTC	CAAATTACTT	127320
ACAGTTATCA	AACCATTACC	CTGCCTAAAT	TGATATTCTA	ATAACAAAAT	TATCTTGAAA	127380
ATTAATTTAA	TTTTCAAGCA	CCATATAAAT	ATATTGACTC	AACTCTCAGT	TTTTTCGTAT	127440
ATTTAATATT	ATTATATAAG	GAGATGTTTG	AGATGAAAAA	TATTAAGCCG	TTAGCTGATA	127500
GAGTTTTAAT	AAAAATCAAA	GAAGCTGAGA	GTAAACAAT	CTCAGGACTT	TACATACCAG	127560
AAAATGCAAA	AGAAAAACA	AATATTGGGA	CAGTTATAGC	TGTTGGTTCT	AACAAAAGAG	127620
AGATCACTGT	AAAAGTTGGT	GATACTGTGC	TTTATGAAAA	ATACGCAGGA	GCTGCTGTAA	127680
AAATCGAGAA	TAAAGAACAT	TTAATACTAA	AAGCAAAAGA	AATAGTTGCA	ATAATAGAAG	127740
AGTAAAAAGC	TAAGTTTAGC	TACTTAGCTT	TAATTTTTAT	TAAATATTTA	ATAAAAATTA	127800
CAAATTTATA	CATAAAACT	TATTATTCTG	ATCAATCAAA	TTAAAAATTT	CAAGCTTACA	127860
AAATTCTGTA	AGCTTGAAAA	AATAAAATTA	AATGAAAAAG	CCAATTTTTA	AAGAAAATAC	127920
CATATATTCA	AGCAAATTCG	ATGACATCTA	TTACAATCCA	AAGCAGGGAA	TTGAAGAGAG	127980
TTTTTATACA	TTTATTAAAG	GTTGCAATTT	AGATTTAGAA	TTAAAAACAA	AAAAAATAT	128040
TTTAATAGCA	GAGTTGGGAT	TTGGAACAGG	ATTAACTTTT	ATATGTCTTT	TAAAATTCAT	128100
AAAAGAAAAC	AACATAACCT	CAAAAATTAA	TTATTATTCT	ATAGAAAAAT	TTCCACTCGA	128160
AAAAAAAACA	ATAATGCAAA	TTTCAAAGTT	CTTTGCTAAA	GAAACCGCTT	ATTTTAAATT	128220
AATGTTGAAA	AATTATTCTA	AAATTCCAAA	AAAAAATTTA	AACTAAAAAA	TAACAGAAAA	128280
TGTTAATTTA	AAAATTTTAA	TTGGAGACGC	CAAAATAAAA	ATCAAAGAAA	TTCCTGAAAA	128340
TGTAGAATAC	TGGTTTTTAG	ACGGATTTAA	TCCCAAAAAA	AATCCTGAAA	TGTGGAGCAA	128400

TGAAATATTT	AATTTAATTT	CTGAGAAAAG	CAGTCCGAAA	TGCAAGCTTT	CAACATTTTC	128460
CTCTGCAAGA	ATTGTAAAAG	ATGGCCTAAA	ACTTGCTAAT	TTTAAATACA	TTCACATAGA	128520
AAAAGGATTT	GGAAATAAAA	GACATATGAT	AAAAGCTCAA	AAAAATTAAA	AATTTATTTT	128580
TAACATAAGT	CGTTAAAAAA	ATCCCAACAA	GTATGATATA	CTTCCAAATG	GCACAAGGAG	128640
AATTTTAATG	ACAAAAAAAT	TGTTTGTGAG	GGTATTAATC	TTTTTAATAT	CCAATAATTA	128700
TGCTTTTGCA	AAAGACACAA	TCAAAGATTT	GTTCTTTATA	CAAGATATAC	TAATAAAAAA	128760
AGAGAAATAT	TCCGAGGTTC	TAAATAATGC	AAGCCTTGAA	GGCATTATTG	AAATTGAACA	128820
TAACGGACCA	TACATTAAAG	ATCACGATTC	AGAAGTTAAA	CTTATCCTAA	AAGAAAACGG	128880
ATATAGAAGA	AATTTCAACT	TTTTTAATCT	TTTAAATACT	AGTAATATAA	TCAAAAGTCT	128940
AAGCTTATTT	GACAGCAGAC	CAAAAAACAT	TAAAGAAAAT	GAAATCATAT	TATTAGAGAC	129000
AAAAATGATT	AAAGAAAATC	CCTATAAACG	ATACAAAGAC	GATGATGATT	TTGAATTAAA	129060
ACTAAGTGTA	ACTCGAAAAA	ATAATCAAAT	TTATTTAATT	CTTGATTTC	ATTCCTATT	129120
TGATCAAAGA	AAAACGTTTC	CATCAATTTA	CATCAAAGAA	GAAGATGTAT	CAACAATAAT	129180
AAACAGCTTC	ATGAACTAC	AAGATTCAAG	CTTTTTATCT	CCTCAAGCTT	CTTAACAATT	129240
AATAGCACAA	AATGTGCTAT	TTCTAATAAA	AAGCAAGCAT	TTTACTGAAA	AGCTAACCAT	129300
AGCCAATTTT	ATTACATAAT	AATTTTTTCAA	TCTTTTTTACA	GATTTTTTTAA	ATTAATAATA	129360
TAATTATTTA	TTTTATTAAT	TAAAGAAGAA	AATTCTACAA	ATTTTAATTT	TTCAGATTCA	129420
ACAATTTCTT	TGGGAGCATT	CATTAAAAAA	TTTTCATTTT	CAAGTTTCTT	TGAAACAGAA	129480
ATATTGAGCA	TTTTTACTTT	TTCAAGCTGC	TTTTCAAGCC	TTATCAACTC	TTTGGTTTTA	129540
TCTATCAATG	ACTTAACATC	TGCATAAATT	TCAAAACCAA	CTGCAGCTAC	ACCAAGCATG	129600
CCATCATAAT	TTTCATTGTA	AAATATATTT	TTAAAATTAA	TCATTCTTTT	TACAATGCTT	129660
TCATTAGCCT	TAAAGTATGC	CTCATATTTA	AAATCAGCAT	CAAACTTCAA	AGCAACATCA	129720
ATTTCAACAC	TAGCAGGTAT	ATTAAATTCA	CTCTTAAGTG	TTCTAATAGC	TATAATAAAA	129780
GTTTTCAATA	CTTTAAAAAT	TTCAAATTCT	TCTTGAAAAT	TATTGGCAAT	ATCAAAATTT	129840
GGATATTCAT	TTAAAGCTAA	AATATCTTCC	TTTTCTGCAA	ATTCAGAATA	AATTTTTTCT	129900
GTAACAAAAG	GAATAAACGG	ATGCAAAATT	AACAATGATT	TTTTAAGAAA	AAATAGCAAC	129960
TTAGAAATAG	CCATATTTTG	AATATCAACA	TTTTCATTAT	TTAAATCAAT	TTTGCTAATT	130020
TCAATATACC	AATCACAAAA	ATCATTCCAA	AAAACTCAT	AAACAAATTT	TGAAGCTTCG	130080
TTATATTTAT	AATTTGCAAA	AGAAGACTCT	ACACCAAGAA	TAGTCGAATT	TAAGCTTGTA	130140

AGCAGCCATT	TGTCAATGTC	TAAATTTTC	AAATCATTTA	ATATTTTCT	TTTTTTTAAA	130200
TTTAAAAGAA	TAAATTTGGA	AGCATTAAAA	ACTTTGTTTG	CAAATTTAGC	CCCAAACATA	130260
AAATCTTTAG	CGTCAATATT	TAAATCTTGA	CCCTGAACAG	ACAAAAAGGA	TAAAGTAAAC	130320
CGCAAAGAAT	CACTTCCATA	CTCATTAATA	ATATCAAGAG	GGTCTATTCC	ATTGCCTAAA	130380
GACTTTGACA	TTTTTTTACC	TTGTTTGTC	CGCAAAAGAG	GTGTTATATA	AACATCTTTG	130440
AAAGGAACTT	GCCCTGTAAA	TTCTAATCCT	GCCATCACCA	TTCTTGCAAC	CCAAAAAAT	130500
ATTATATCGT	AAGCTGTTAT	CAAGGTATTT	GTGAGATAAT	AATTTTAAA	ATCAACATCA	130560
ACATTGGGCC	ATCCAAGCGA	AGAAAAGGGC	CATAGCCAAG	AAGAAAACCA	AGTATCAAGA	130620
ACATCTGGAT	CTTGAACAAA	CCTCTTCCCC	ATATCTTTT	CATCTAAAGA	AGGATCAGTA	130680
TCACTAACAA	TAAGTTCAGA	TGTATCAACA	TTGTACCAA	CCGGTATTCT	ATGTCCCCAA	130740
ACAAGCTGTC	TTGATATACA	CCAATCTCTA	ATATTTGATA	ACCAATATTT	ATATGTATTT	130800
TCCCACTTTT	TAGGATAAAA	TTTAAATTCG	CCATTCTCTA	AAGCCTTTAA	AGCTTTGTCT	130860
GCTAAAGGCT	TCATTCTCAC	AAACCACTGA	GTAGACAAAT	AAGGTTCAAT	AACCTCACCT	130920
GACCGATAAC	AATGCCCAAC	CTGTTGTTTA	TGCTTCTTAA	CATCTTGCAA	AAAACCCTTT	130980
TCCATTAATT	CTGTTTCAAT	TTTAAATCTT	GCATCTTTCG	CACTTAATCC	TTGGTATTGC	131040
AAAGGAACAT	TTTTATTAAG	TTTTCCATCT	TGAGTTAAAA	TATTGACCTT	AGAAATATTG	131100
TGCCTTTTTG	AAATTTCAAA	ATCATTAGGA	TCGTGTGCAG	GAGTAACTTT	TAAAGCCCCA	131160
GTGCCAAAAG	CGCTGTCAAC	ATAAAAATCT	GCAATAACTT	TTATCTTTTT	AGTTGTCAAA	131220
GGAATTGTAA	CTTCTTTGCC	AACTAAAGAC	TTATATCTCT	CATCATTAGG	ATTAACAGCA	131280
ATAGCAGTAT	CCCCAAACAT	TGTCTCAGGC	CTAGTTGTTG	CAACCTCAAT	AAAAGAAGAG	131340
TTATCAATAA	AATACTTAAC	AAAATAAAGC	TTACCATCAA	CTTCTTTGTA	TTCAATCTCT	131400
TCATCGCTAA	CAACACTCCC	AGATCCAGGA	TCAAGATTAA	CAAGATACTC	ACCCCTATAA	131460
ATCAACCCCT	TAAAATACAA	GTCCTTAAAA	ACCTTGTTAA	CAGCCTTACA	AAGATTCTCA	131520
TCAAGAGTAA	ACCTTTCTCT	TGAGTGATCA	TAAGACGCCC	CAAGTTTGTT	TATCTGATTA	131580
ACAATTATCC	CTCTATGCCT	ATCTTTTAAT	TTAAAAATTT	CTTGAACAAG	CTCTTCTCTT	131640
TCAAAATCAT	CTTTGCTTTT	ACCAATCTTT	TTAAGATGTC	TTTCAAAAAC	AGCCTGCGTT	131700
GCTATTCCTG	CATGATCTGT	GCCAAAAAGC	CACAAAGTAT	TGTGTCTTTT	CATTCTTTTA	131760
TACCTTACAA	GAACATCTTG	CAAAACAAAA	TTAAGAGCAT	GCCCCATGTG	CAACACGCCA	131820
GTAACATTAG	GAGGAGGCGC	AACCATACTA	AATTTTTCAT	ATAAAGAATT	ATCTGGCAAA	131880
AAAACATTGT	TTTTAAGCCA	CTTAGTGTA	ATTTTCATCT	CAAATGCCTT	AGGATCATAT	131940

TTTTCAAGAG	GTCTACAATT	CATCGTTTAA	AAATCTCCTA	CCATGTCAAT	AACTTTTAAA	132000
TTTTCAAAAA	GCTCTTCATA	AGTTTTTGAA	TTTATAAGAT	TTGAGAAATA	CTTACTCTGT	132060
CTTTCCTTCT	CAGTAAAATA	TATCTTAAAA	TATTTTTTAA	GTTTATTAAA	ATCTTTTTCA	132120
AGTCCCCAAG	TAGCATGAAA	ATCCTTTATA	TGGAATTTTA	ATATATTTAA	CCTAAAATTT	132180
AAATTACTAT	CTAAAAAGTG	TTTTGAAAAG	GGCTTAAATA	AATTCAAATT	ATCAAAAATT	132240
CCACGACCAA	ACATAATTCC	ATCAATTAAA	TATTTATCAA	CATAATATCT	AGCTTCCTTG	132300
AGACTCAAAA	CATCCCCATT	TCCAATAATC	AGAGTAGAAG	GACTAAGATT	ATTTCTGTAAT	132360
TTAACAAGTT	CATAAAAAAT	ATCAAAATTA	ACAGGACCTT	TGCTCTGATT	AACAGCAAGT	132420
CTTGGATAAA	CCGTTAACAT	ATCAATTTCA	AGGCCTAACA	AAAACCCCAG	CCAATCATCA	132480
ACTTCTGGAT	ATGAAAATCC	ATGCCTGGTC	TTAACACTAA	GAGGCAAATT	AAATCTCGCA	132540
CAAGCTTCTT	TGCTTGCAAG	AATTATCTCT	TTAGCTAAAG	ATTTATTATT	AATTAAAGCT	132600
GAACAAACTC	CTTTTTTAAT	TATTTTACTC	TTAGGACAAC	CCATATTAAT	ATCAATTCCC	132660
CAAAACCCCA	TGCTCCCTAA	TATTTCTATT	GCTCTATAAA	ATTGCTCAGG	AACATTGCCC	132720
CAAATCTGAG	CAATTAAAGG	CCTATTAAGT	TCATTGGGTT	TTAAAAAAC	ATGTTGAACA	132780
GATTGTTTTG	ATCCATTTAA	AATTCCTTTT	GTAGAAATAA	ATTCGGTAAA	ATAAATATGA	132840
GGTTCTCCTT	CTGCAGATCC	TATTAAATGA	ATTAAATTTT	TAAAAACAGT	ATCGGTAACA	132900
TCTTCCATTG	GAGCTAAAAT	CATAATTGGA	AGAGGAATAT	CAAATAAAAA	CTTCATAAAA	132960
TAAC TACATA	ACAATAAAAT	ACAAATTTAA	ACAAAGTAAA	ACAGTCGTTT	AATAAATTTA	133020
AATTTCCAAA	AATAACATTG	AAATTTTTTAA	AATAAAGAAA	TACCACAAAA	GGACCGGCAA	133080
TAAAGGACAT	AATCATAAAA	TTTTAAAAAG	TTAATAAATA	ATTTTAATAA	AAAAATAATT	133140
TAAAAAATCA	ATCCATAATA	AGGCTCTCGC	CTGGAACCTT	GGTTTTTCTT	ATTAAAGTAT	133200
CTTTATATCC	AGCAGATTTA	ATAAGTTCTA	TATTTTTTTT	AACATCATCG	GCATTAGTAG	133260
GAATAAAAAAC	GGTATAAAAC	GGCCCATGAG	AATTTACTAC	AACAAAAAGA	CCCGCTTTTT	133320
TTAATATTCG	ATAAGCCCTA	TCAGCGTAAT	CTTTTTTCCT	ATAAGATCCA	ACTTGTATGT	133380
AAAAATCTGT	TTCTTTGTCA	GCAGAACTG	AGTAATCTGC	CAATAAATCT	TTGGATTTAT	133440
TTTAAACAGA	AATATTTTCG	TCTGATTGCT	TGGTTTTCTT	TTCTTCTTTA	ACCCAGGAA	133500
TATTAAATGT	TTTTTTTAAAG	TCACTAGATT	TTGATACAGA	AGGTCTTTTT	TCATTCTGAAC	133560
TTTCAATCAC	TTCAATTTTT	ACAGGAGCAA	CCCCTATTCC	TAAAAAATCA	AGCTTCTCAG	133620
CGGCATATTT	TGACAAATCG	ATTATTCTAT	CCTTCCTAAA	AGGACCTCTA	TCATTAATTC	133680

TTACAACAAC	TGATCTATTA	TTTAAAGAT	TTGTAACCTT	TACGGTAGTA	TTTAAAGGGCA	133740
ATTCTTTGTG	AGCAGCAGTA	AGCGCCATCA	TATCAAATTT	TTCGCCATTA	GCAGTAGTTT	133800
TGCCGTGAAA	AGCTTCGCCA	TACCATGAAG	CAAGACCCAC	TGTGGCAGAA	TTTAAATGAG	133860
AAGCAATAAA	AAAAAATACA	AAGAGAAAAA	CAAAGTTTTT	ATTATCTCTT	AAGATGGCAT	133920
CAATTAAATT	TCTCATAATG	TTTATTATAA	TATAAAAAACA	TATTTCAATA	AACAATTAAG	133980
CTTGCAAATT	GCTTATTTTAC	ATTTTTTTTTG	ATTTAATTAT	AAAAAGAAAA	AAGTCTAAAA	134040
AATGATATCA	ACAGAAATAA	TTAGCAGCAG	CCAAATACAA	AAAGCAGCAA	AACTTATCAA	134100
AATGGGAGAA	CTTGTAAGTAT	TCCCAACAGA	AACAGTTTAC	GGAATTGGCG	CAAATGCTTA	134160
CAATGAAGAT	GCTGTAAAAA	TGATTTTTTTT	AGTAAAAAAA	AGGCCCATCA	ACAATCCTTT	134220
AATAGTACAT	GTTGATACGG	TAAAAAAAAT	AAAAGAATTA	TCAGAATATA	TTCCCAAAAG	134280
TGCCCTCATG	CTAATCAAAA	AATTTAGTCC	AGGCCCTTTA	ACTTATGTTC	TTAAAAAATC	134340
AATAAAAATA	TCTAGATTTG	TAAGTGGAAG	CCTAGACACA	GTGGCAATAA	GAATTCCTGC	134400
AAATAAAACA	GCTTTAAGCC	TAATAAAAGC	ATCTAAAGTC	CCCATAGTAG	CACCGTCTGC	134460
AAACATATCA	AAAAGACCAA	GCTCAACAAA	TTTCGAAATG	GCCTTAAAAAG	AATTAAATGG	134520
ACTTGTAAGA	GGAATAATAA	AACCGGAAGA	GAACAAAGAC	TTTAATATTG	GAATCGAATC	134580
AACTGTGGTT	GGGTTTGACC	TAAAAGATAA	CGTACTGATA	TTAAGACCAG	GCGCAATAAC	134640
AAAAAAAATG	ATAGAAAATG	AACTTCAAGG	AAAATATACA	GTAAATTACG	CAGAAACAAA	134700
AATGGAAC TA	GAAAAATCAC	CTGGAAACAT	AATTGAACAT	TATAAGCCAA	AAATTCCCGT	134760
TTATTTATTT	AAAAGTCAAG	ATAACATAAG	AAGATACTTA	AACAAAGATA	CGAAAAATACT	134820
TATCACAAAA	GCTACTCTAA	AATCCTATTT	ATTCAATTTT	TTTTGGAATA	AAAAAAATAT	134880
TACAGTATTT	AACACTCTTG	AAGAATATGC	ACAAAACCTT	TACAAAGAGT	TGGTAAATTC	134940
TGAAAACAAC	TACAAACAAA	TACTTAGCGA	ATTCTTAAAA	GACGAAGAAC	TTGGACATTC	135000
AATAAACAAAT	AGAATCAAAA	AAGCTAGTTC	AAATAGATTC	ATTAACAAAA	AATGACGCTA	135060
AATTGTTATT	TAAAAATAAT	CAAAAAGCAT	AAATATTCAT	TAATAAAATA	ATGCTAAAGC	135120
TAAAAGCAAA	AGTCTAAAAC	ACGCCACACC	TCTCCTCCCA	ACCCGAGCAA	AACCAGCAAA	135180
TACAACCAAG	AAGACTTAAA	CTTAAATATT	ACAAGTAAAT	TTTCGCATAA	TCACATAAGA	135240
AAAATTTCAA	TCCTTTGATT	AATTGAAATA	ATCATGGATC	AACATAGTAT	ACTCAAGTGG	135300
TATTTTATCT	TCATCAAAAAG	CAATACCAAG	CGCAAAAACC	TTTCCGCTGG	GAGTCTGAAT	135360
AACGCTTAAG	CTTTTTGATT	TACCTTCAAT	AAAAATTTCT	CCATCTATAA	ATTCAAAACT	135420
AAAAATCAAA	TCAATTGCAT	CTTCCTCGAC	ATCCCCATAA	TCAAAAGAAG	AAATTACTAA	135480

AGCACCCCCA	TAAGATAAAAT	CTTTTATTAA	GCATTTATGT	TTTGCTCCAT	TAAACTTGAT	135540
AAAAGCTTTA	TCAGAATCAA	TTTTTAGCTT	TCTAATAGAA	TCTTTATCGA	TAATAATCCT	135600
CTCATGAATT	CTCTGATTTT	GCCCAAGCTT	TAAATCAAGA	AGCTTTCCAA	CTTTAATAGC	135660
AATCTCTTCT	GGTGCAGGAG	ATAAAAAATC	TAATGTTAAT	AAATTGTATT	CTTTATTCAA	135720
AGAAGAATAA	GCAGAAGCAC	TCAATAGTTT	TACAGACAAA	AAAGGGAAAA	AAGCTGCACT	135780
ACTCTTAGAA	TCTGAATTTT	TCTTAAGTTG	AATAGAGCCT	AAATTTTTAT	TTTTAGCCAA	135840
AGCGGGCAAT	ACTGTATCTT	CTTGAAAAAT	AAGCTTAAGA	GAATCCATAG	AAATAGAATA	135900
AATTACTCCA	AAAGCGGTAT	AAGAGCCTAT	TCTCATCTCA	ATAGTATTGC	GAAGATTTAA	135960
AAAACATTTT	ATCTCTGTGC	TCATTTTAAT	CTCTTTACCC	CTATACTTAG	CCCCATAATC	136020
TCTTATTTTT	CTAGATAAAA	GCATAAACCT	CTCCTTTCTC	CTTTTTTTGA	ATATAAAAAT	136080
ACTAAACAAA	GCTTCAAATT	GTTTTAAACA	GTTTTACACA	AATAAAAAAG	TTTAAACAA	136140
TTAAGCTATA	AACACGCTTA	GCCAAAAATC	AATAAATCTA	CCCTTAAAT	CAAAAACATA	136200
CCATCATCCA	CCTCTCACAG	CATCTTTCAA	TTAACAAAAC	TTACAAAATG	CTGTTTACAT	136260
AGTGTAAT	TATAACAATA	ATTTTACACT	ATAACAATCA	ACCCATAACA	TTATTAATTC	136320
TTATGCACTT	ATAGTATACT	TTAGTAAAAA	GTATGAAATT	GAATAGCCCT	AATTTGAAAA	136380
AAATAAATAC	GCATAAGCTG	CTTATATATT	TAACATATTT	CGCAGTTAGC	TTTTCTATTA	136440
TCACACTCTC	ATTAGCAGTA	TCTAAGACTA	TAAACATACA	AAAAGATAAA	AATTTTCGGAT	136500
ATGTAAATCC	AGCAGTTCCT	TCAAGACTTT	TAGATATTAA	TGGAAAACAA	ATAACTCAAT	136560
TTATATCTGA	TGAGAACAGA	GAATTAATGC	CTTTGAGAAA	AATGCCTGAC	AATCTAATTA	136620
ATACGCTTTT	GATACGGGAA	GATATTGGTT	TTTTTCTCA	TCGAGGTTTT	TCCTTGATAG	136680
GAATATTTAG	AGCCGCATTT	AATATTGTTC	TTGGCAGATA	TTTTTCAGGC	GGCAGCACAT	136740
TAACCCAACA	ACTTGCAAAG	CTTCTCTACA	CAAATCAAGC	AAGAAGATCT	ATTTTGAGAA	136800
AATTACATGA	AATATGGTGG	GCAATTCAAC	TTGAAAAAAA	ACTCTCAAAA	TACGAAATAC	136860
TAGAGAAGTA	CCTTAATAAA	GTTTATTTTG	GAAACGGAAA	CTATGGAATA	GTTGCAGCAT	136920
CAAATTTCTT	TTTTGGCAAA	AGTGTAATA	AAATCAATAC	AGCAGAATCA	GTAATGATGA	136980
TAATCCAGCT	TCCAAATGCA	AACTTTTATT	CACCTCTTTA	CAATCCAGAA	TTTTCAAAAA	137040
AAATACAACG	TGCAGTTTTA	AACCAAGTTG	TATCAAATGG	AATAGTCAAG	GCTGAAATTG	137100
CTGAAAAAGA	ATTTAATGAA	TACTGGCAAA	ATTATGATTG	GACTAGAATG	GCTGACACAT	137160
CTGCAATTTT	AAACAAAAAA	GACCAAGCTC	CTTATTTCTC	TGAATATATA	AGGCAAAAAA	137220

TACTAAAATA	TTTACCAGAT	GGCGCAAACA	TATATAAAGA	TGGGTACTCA	TTATATTCAA	137280
CCCTTGATCT	TGAAGCACAA	AAATATGCAG	ATAAAGTTAC	AAACGACATG	ATTAATAAAG	137340
CAAGAACAAT	GCACAATTTA	AATAGATCAT	CTGAAACAAT	AATCATTAAAT	TCAGAAAATTG	137400
TCCCTGTTAGT	AGATGCGATA	TCAGATTTAT	TGGGAATTAA	AAATTTAAGA	ATAAAATGGAA	137460
GACAATATAA	AAAAGTGAAG	AAAAGAAAAT	TTTACGAAGA	CAATATTGAT	CTAATTGCAA	137520
GTTTTGGAGC	TATACTTGGA	ATTGATAAAA	TAGATAAGGC	GACAAAAGAA	TATATTATCA	137580
AAAATAAATT	AACACCGAAA	CTTATTGCAC	AGCCTGAAGG	AGCAATGATA	GCAATAGATA	137640
CAACAAGTGG	AGCAATAAGA	GCCATGGTTG	GGGGAAGTGG	ACACACTAAA	GACAATGAAT	137700
TTAATCGAGC	CACACAAGCA	AAAGTTCAGC	CTGGAAGTGC	ATTCAAAGCA	TTATATTTTG	137760
CAGCCGCAAT	TGATCTAAAA	AAAATAACAG	CTGCGACAAT	GTTTTTCAGAC	TCTCCAGTAG	137820
CATTTCTAAA	TAAAAATGGA	GAAGTTTATG	CTCCGGGAAA	TTATGGCGGC	AAATGGAGAG	137880
GCAACGTTTT	AACGCGCCAA	GCATTAGCTT	TGTCCTTAAA	TATTCGGGCA	TTAAGAATAT	137940
TAGACCGGCT	AGGCTTTGAC	TCTGCAATTA	GCTACTCCTC	AAAAC TACTA	GGAATAACAG	138000
ATCCAAAAGA	AATAGAAAAA	ACGTTTCCAA	AAGTTTATCC	ACTAGCGCTA	GGTGTAATAT	138060
CAGTTTCTCC	AATCCAAATG	GCAAGAGCCT	TTGCAATTTT	AGGAAATAGT	GGTAGCGAAA	138120
TCGAACCTTA	TGGGATAAGA	TACATTGAAG	ACAGAGCTGG	AAGAATAATA	ACAAATGAAG	138180
AAGCAAGCAT	ATTGGCTAAA	ATAAAAAACA	AAGAACACCA	AACTCAAATA	GTATCTCCTC	138240
AAACCGCTTA	CATAATCACA	GATATGATGA	AATCAACAAT	TCAATACGGA	ACCCTAGCAA	138300
ATCAAAGATA	TACAAATCTC	AAAAATTTTA	AATCAGACAT	TGCTGGAAAA	TCGGGAACAA	138360
CACAAAATTG	GGCAGACGGA	TGGGCAATAG	GATACTCTCC	TTATATAACA	ACAGCATTTT	138420
GGGTTGGATT	TGACAAAAAA	GGATATTAC	TGGGAATATC	TGGAACAGGA	ACAGGATTGG	138480
CAGGGCCTAG	TTGGGGAGAA	TTTATGGCAG	AATATCACAA	AAACTTACCC	AAAAAAGTTT	138540
TTGTAAAAACC	TGCAGGAATA	ATTAGCATCC	CCGTACAAGC	AGAAACGGGT	CTACTACCGG	138600
AAGAAATTGC	TGATGAAAAA	ATAATAAATG	AACTATTTAT	TTCCGGCACC	CAGCCAGTTG	138660
AAAAATCAAA	ATATTATGAA	AATAACAAG	AATTTAAAAA	TACAATAGAA	TTTAACATAT	138720
ATGGAATTGA	TGAGATTAAT	AATAACGATG	AAATAAATTT	TGACACTCCT	GAATTTGAAT	138780
ATCTTGATAA	TAATCTTGAA	AGCTTTAATA	ACAATAGTAA	TAATGATAAT	AATCTTGAAA	138840
GCTTTAACAA	TAATAACAAT	GATCTTGAAA	GCATTAATGA	TAATGAAGAA	AATAAAAATG	138900
AAGATGAAAT	AGAAATGAAC	ATTGAAGAAC	CCTTAAATGA	AATAGAAAAT	AAAAATCCAC	138960
AACAAGATCT	AGTTAATAAC	AATAATAACC	AGGAAATGCT	TATTGAAAAC	ACCAAAGAAA	139020

TTAAAGACGA	AGTCATTGTT	AATGAAACAA	ACATAGAAAC	ACAAAGCACA	AAAGAATTAA	139080
ATTCAAACAA	CAATGAAAAT	GAAAAAATTA	ACAACAAAGA	CGTCAACGGA	GAAGATATCC	139140
AATTGGATTA	AAACAATATG	TTAATAGATA	TCGATCAAAT	AAAAATAAAA	AAAAGAATTA	139200
GAAAAAATAT	AGGAGACATT	GAAACTCTTA	AAAACAGTAT	TATAAAACAT	GGATTAATTT	139260
ATCCAATAAT	AATAGATAAA	AATAAAAACT	TGATAGCAGG	ACTTAGAAGA	TATCAGGCCT	139320
TAAAAGAAAT	AGGCTATAAA	GAAATTGAAG	TAAAGGTAAT	CTCAATTGAA	AACAAAAAAA	139380
CTTTACTTGA	AATTGAACTT	GATGAGAATA	ATGTTAGAAA	ATCATTCAAC	AGAAGCGAGG	139440
CAAACGAAGG	AGAAGCTTAC	TTAAAAATTT	ATTCTGAAAG	CAATATAATA	ATAAGATTCC	139500
TTAAATTTAT	TATCTTAAAA	ATTAAAAACA	TGTGTAAAT	AAGAAATAGA	AAAATTTAAA	139560
TCATAATAAA	GAGGTGTGTT	TATGTTAAAT	TACAAAAATC	TTAATGAACT	TGAAAATTTT	139620
AAAATCCTTG	AAGGTATTGC	TCCAGAAGTG	CTCAAAACGG	CATTAAGTGG	AAAAAGGATA	139680
AAAGAATACG	ACATTACAAT	AGAAGGAGAT	AGTGTACATT	ATAACTATGC	TTCAAAACAA	139740
ATTAATGAAA	CCCACCTTAA	AATTTTTCAT	AATTTAAGCG	ATGAAGCAAA	TTTAATAGAA	139800
AAATATAAAG	AAGTGCTTGA	TGGGGAAAAG	ATCAATATTA	GTGAAAATAG	AAAAGTCCTG	139860
CATCACCTTA	CAAGAGGGCA	AATTGGTAAG	GACGTAATAG	AAGACAATAA	AGAAAAATATG	139920
AGAGAGTTTT	TCCAATCAGA	ACTTGAAAAA	ATATATAATT	TTGCAAAGCA	AATTCATTCT	139980
GGGAACATTA	AAAGTTCAAA	TGGCAAAAAG	TTTAAAAATG	TAGTTCAAAT	AGGAATTGGT	140040
GGATCTAGCC	TGGGGCCAAA	AGCTCTTTAC	AGCTCAATAA	AAAATTATGC	AAAAAAACAC	140100
AATCTAGCCC	TAATGAATGG	TTATTTTATT	TCAAACATTG	ATCCAGACGA	ATCAGAAGAA	140160
GTATTAAGCA	GCATTAATGT	TGATGAAACG	CTTTTATTAT	TTGTCTCAAA	AAGTGGAAAT	140220
ACATTAGAAA	CTAAAGCTAA	TATGCAATTC	TTAATAAACA	AATTAAAATT	AAATGGCATA	140280
AAAGAATATA	AAAAACAAAT	GGTCATTATA	ACACTAAAAG	ATAGCATGTT	GGCAATAGAA	140340
GAAAAAGGAT	ATCTTGAATA	TTTCTTCATG	CATGACTCAA	TAGGTGGAAG	ATTTTCTCCA	140400
ACATCAGCAG	TTGGACTTAC	ACTACTTACT	CTTTGCTTCA	CAGAAAAAGT	TGCAAAAAGAA	140460
ATTCTAAAAG	GAGCCAATGA	GGCTGACAAA	AAATCATTTA	ACAAAAACGT	AAAAGACAAT	140520
GCATCTCTCT	TGGCAGCACT	AATTAGCATA	TATGAAAGAA	ATGTTCTAAA	TTACAGTAGC	140580
AACTGCATCA	TTGCTTATTC	TAAAGCAATG	GAAAATTTTT	ATCTTCATTT	ACAACAACCTT	140640
GAAATGGAGA	GTAATGGAAA	AAGTGTAAC	AGATTTAATG	AAACAATAAA	CTACAAAACT	140700
GTAAGAATAA	TTTGGGGAGG	CATTGGAACA	GATGTTCAAC	ACTCATTCTT	TCAAATGCTT	140760

CACCAAGGAA	CGGATATAGT	CCAATGGAT	TTCATAGGTT	TTAATGAAAC	CAAACTTAAA	140820
GAAGATGTAA	TATCTGATAA	CAGCTCAAGC	AATGATAAAT	TAAAAGCAAA	TTTAATAGCC	140880
CAAATAATAG	CATTTTCAAA	AGGTAAAGAA	AATAGCAATA	AAAATAAAAA	TTTCCAAGGC	140940
GAGAGACCTT	CTGCACTAAT	ATATTCAAAA	GAATTAACAC	CTTATGCAAT	AGGAGCAATA	141000
CTCTCCCATT	ATGAAAATAA	AGTAATGTTT	GAGGGATTTT	TATTAAATAT	AAACTCATTC	141060
GACCAAGAAG	GAGTTCAGCT	AGGAAAAATT	ATTGCAAATC	AAATTTTAAA	AAATGACAAT	141120
TTTAAAGATG	AAGTAATAGA	ATCTTATTCT	AAAAAAATTC	TAAAAAAATT	TTAAAACAAG	141180
ATTAATTAAT	TTTTGAATAT	ACCCCCTTAA	GTTTAAAAAA	GAATGCACTA	AGCTTATATA	141240
AGAGGTAATA	ATGGATAAAA	TAAGTATATT	ATATACATTA	ATCAATATTA	TAATAATGCT	141300
TATTCTAATA	AGCATAGTTT	ATCTTTGTAA	AAGAAAAAAT	GTTTCTTTTA	CAAAAAGAGT	141360
GTTTATAGCG	TTAGCAATCG	GAATAGTATT	TGGAATGACC	ATTCAATATT	TTTATGGAAC	141420
AAATTCAGAA	ATAACAAACG	AAACTATAAA	TTGGATAAGT	ATTTTGGGCG	ATGGATACGT	141480
AAGGCTCCTT	AAAATGATTA	TAATCCCCTT	AATAATAACA	TCAATAATCT	CTGCAATAAT	141540
AAAATAACC	AATAGTAAAG	ATGTTGGGAA	AATGAGCCTA	CTTGTAATAT	TAACACTAGT	141600
ATTTACAGCA	GGTATTGCTG	CCATAATTGG	CATTTTCACT	GCTTTAGCAT	TGGGATTAAC	141660
AGCCGAAGGA	CTACAAGCGG	GAACCATCGA	AATTTTACAA	AGTGAAAAAT	TGCAAAAAGG	141720
CCTTGAAATA	TTAAATCAAA	CAACAATCAC	AAAAAAATC	ACAGATCTTA	TTCCACAAAA	141780
TATATTTGAA	GATTTTGCAG	GGCTTAGAAA	AAACTCAACC	ATCGGGGTCG	TGATATTTTC	141840
AGCTATCATA	GGAATAGCCG	CCCTTAAAC	ATCTATCAAA	AAGCCAGAAT	CAATAGAATT	141900
TTTTAAAAAA	ATAATATTAA	CACTCCAAGA	CATAATATTA	GGTGTAGTAA	CTTTGATTTT	141960
AAAATAACG	CCTTATGCTA	TATTAGCTTT	AATGACAAAA	ATTACAGCAA	CCAGCGAAAT	142020
CAAAAGCATA	ATAAAGCTTG	GAGAATTTGT	AATTGCTTCC	TACATTGCCA	TAGGTCTTAC	142080
ATTTCTTATG	CATATGACAT	TAATTGCAAT	AAATAAATTA	AACCCAATTA	CTTTTATAAA	142140
AAAAATATTC	CCAGCACTAT	CATTGTCATT	CATATCTAGG	TCGAGTGCTG	CAACCATACC	142200
CATTAATATA	GAAATTCAAA	CTAAAAATCT	GGGAGTAAGC	GAAGGAATAG	CAAATTTATC	142260
AAGCTCCTTT	GGAACATCAA	TTGGGCAAAA	TGGTTGTGCA	GCACTACACC	CCGCTATGCT	142320
TGCAATAATG	ATAGCACCAA	CTCAGGGAAT	AAACCCACA	GATATTTTAT	TTATACTCAC	142380
ACTTATTGGA	TTAATAATAA	TAACTTCATT	TGGAGCTGCT	GGCGCTGGTG	GAGGCGCAAC	142440
AACAGCCTCA	CTAATGGTGC	TCTCAGCAAT	GAACTTTCCA	GTGGGATTGG	TAGGACTTGT	142500
AATATCTGTT	GAGCCTATAA	TTGACATGGG	AAGAACAGCT	GTTAATGTAG	GCGGCTCAAT	142560

GCTTGCAGGC	GTTATATCTG	CTAAACAGCT	CAAACAATTC	AACCATAATA	TATACAACCA	142620
AAAAGAGCTT	GTAAACAAAT	AAATAGGAAA	ACAATGATGA	AAATAATAAT	TATTGGGGGC	142680
ACATCAGCAG	GAAGTAGTGC	CGCAGCTAAA	GCAAACCGCT	TAAACAAAAA	GCTAGACATT	142740
ACTATCTATG	AAAAACAAA	TATTGTATCT	TTTGGAACCT	GTGGCCTGCC	T'ACTTTGTG	142800
GGGGGATTCT	TTGACAACCC	CAATACAATG	ATCTCAAGAA	CACAAGAAGA	ATTCGAAAAA	142860
ACTGGAATCT	CTGTTAAAAC	TAACCACGAA	GTTATCAAAG	TAGATGCAAA	AAACAATACA	142920
ATTGTAATAA	AAAATCAAAA	AACAGGAACC	ATTTTAAACA	ATACTTACGA	TCAACTTATG	142980
ATAGCAACTG	GTGCAAAACC	TATTATTCCA	CCAATCAATA	ATATCAATCT	AGAAAATTTT	143040
CATACTCTGA	AAAATTTAGA	AGACGGTCAA	AAAATAAAAA	AATTAATGGA	TAGAGAAGAG	143100
ATTAAAAATA	TAGTGATAAT	TGGTGGTGGA	TACATTGGAA	TTGAAATGGT	AGAAGCAGCA	143160
AAAAATAAAA	GAAAAATGT	AAGATTAATT	CAACTAGATA	AGCACATACT	CATAGATTCC	143220
TTTGACGAAG	AAATAGTCAC	AATAATGGAA	GAAGAACTAA	CAAAAAAGGG	GGTTAATCTT	143280
CATACAAATG	AGTTTGTAAG	AAGTTTAATA	GGAGAAAAAA	AGGCAGAAGG	AGTAGTAACA	143340
AACAAAAATA	CTTATCAAGC	TGACGCTGTT	ATACTTGCTA	CCGGAATAAA	ACCTGACACT	143400
GAATTTTTAG	AAAACCAGCT	TAAACTACT	AAAAATGGAG	CAATAATTGT	AAATGAGTAT	143460
GGCGAAACTA	GCATAAAAAA	TATTTTTTCT	GCAGGAGATT	GTGCAACTAT	TTATAATATA	143520
GTAAGTAAAA	AAAATGAATA	CATACCCTTG	GCAACAACAG	CCAACAAACT	TGGAAGAATA	143580
GTTGGTGAAA	ATTTAGCTGG	GAATCATACA	GCATTTAAAG	GCACATTGGG	CTCAGCTTCA	143640
ATTAAAATAC	TATCTTTAGA	AGCTGCAAGA	ACAGGACTTA	CAGAAAAAGA	TGCAAAAAAG	143700
CTCCAAATAA	AATATAAAAC	GATTTTTGTA	AAGGACAAAA	ATCATACAAA	TTATTATCCA	143760
GGCCAAGAAG	ATCTTTATAT	TAAATTAATT	TATGAGGAAA	ATACCAAAAT	AATCCTTGGG	143820
GCACAAGCAA	TAGGAAAAAA	TGGAGCCGTA	ATAAGAATTC	ATGCTTTATC	AATTGCAATC	143880
TATTCAAAAC	TTACAACAAA	AGAGCTAGGG	ATGATGGATT	TCTCATATTC	CCCACCCTTC	143940
TCAAGAACTT	GGGATATATT	AAATATTGCT	GGCAATGCTG	CCAAATAGAA	AGAATTAAAT	144000
TAATTTAATT	CTTCATGCTA	ATTGGTTGCC	CCGTACTTGA	AAGAACATCT	CTCCAAAAAG	144060
AACCATTTGG	ATTAACCTTA	TTTCTGTCAA	TTACTGCCAT	CTTAATAGGT	ATATGAACAA	144120
ATTTTGTAAT	CCATAAACTA	ATCAACATTT	TTGTCTTACC	AGCCATTGCA	GCATGCACAG	144180
CATTCGACCC	AAGCCTAGCA	CAATAAAGCG	AATCACTGGC	ATTAGCAGGT	GAACCTCTAA	144240
TAATATAGCT	GGGATCAATG	TATTTAAGAG	TAAATTGTAT	ATTTTTTGCT	TTAAAATATT	144300

CTGTAATTTT	ATCTTTAATA	TAAAGCCCAA	TATCCTCATA	AAGCAAATTC	CCAGAATCGT	144360
CTTTCTTCTT	AGGAAAATGA	TCAAAATATT	TTTGGCCTGC	TCCTTCTGCT	ATCAATATTA	144420
CTGCATGGGG	AATCTCTTCT	AAGCTTTCTT	TCTCTAAAAG	TCGTCTTTCA	AGATGAACAA	144480
GAAATCCATT	AGGACCTTCT	ATGTCAAAAT	CAAGTTCTGG	GATTAAACAA	AAATTAACAT	144540
CATTAGAAGA	AAGTGCGGTA	TGAGCAGCAA	TAAAGCCAGA	ATCCCGTCCC	ATAACTTTAA	144600
CAAGTCCAAT	GCCATTATAA	GCACTATTAG	CTTCAAAATG	AGCACCAGCA	ACAGCTGCAA	144660
CAGCTTGTTT	TACAGCAGTC	TCAAATCCAA	AAGATTTTTG	AACAAACATA	AAATCATTGT	144720
CTACGGTTTT	AGGAATGCCC	ACAACCTGCTA	TTTTTAAATT	TCTTTTTTCT	ATCTCCTCAG	144780
CAATAAGAAG	AGACCCCTTT	TGAGTACCAT	CCCCGCCAAT	GTTAAAAATC	ATATTAATGT	144840
TCATTCTCTC	TAAAGTATCA	ACTATTTCCA	CAGGCTTAAT	ACCACCCCTT	GAAGAACCAA	144900
GAATAGTACC	TCCAAATTTA	TTAATATCAT	CAACAACATC	TGGATTAAGA	TTAATAAAAG	144960
GTGAATTTGA	CTCAGGAAGA	AGCCCTTGAT	ATCCAAATTT	TACTCCATAA	ATATTGCGAA	145020
CCCCATATAT	TTTCCATAAA	GTTTCGCACAA	TAGAGCGAAT	AACATCGTTA	AAACCAGGAC	145080
AAAGCCCACC	ACAAGTAGTA	ATAGCAGCTT	TAACATGCCT	GGGCACAAAA	TAAATTTTTT	145140
CTCTAGGCCC	AGCTTTTTCT	AAAAGAACAT	CTTCATACCT	ATCTCCCTTA	TCCTCATTC	145200
TATATACACT	AAACTTGATT	TTATTTTTTT	CATTAACAAA	ATGGGAAGAA	CCCTCACTAG	145260
CATAAAAATC	AATCAAAGGA	TTGTTTTGCT	TGCATTCTCC	CAAGCTATCT	ATTTTAAAT	145320
CTAAATTTTC	ATTTTTAATT	CTATACACCA	AATACTCCTT	TATAGAATTA	TAACCTAATT	145380
ATTTTCTAAT	AAATCGACTT	TGATCTTTAA	TCATATCGTA	TATGTCATCG	TAAATATAAG	145440
GAGACCCCTC	AATAGGAGAT	TTAATTAAAT	TACCAGCTAT	GAATTCAAAA	TATTTATTCA	145500
ACTTTGAATT	TTTCTCAAAA	TCAATAAATG	GAACTCTATT	ATTAATAGCC	TCTCTGAAAC	145560
TTTTTGCAAA	AGGCACAAAA	CCTATAAACT	CTATTGGTAT	ATTAATATTA	TTCTTAACAA	145620
CATTAATCAA	ATTTTCACAC	ATAGCAATCT	CTTCACTAGT	TTCTATTCTA	TTTAGCACCA	145680
CTCTAGGATA	AAAATTATTC	ATCATCCTCT	TAACTTTCAA	AGAGGAACTC	AAAGAAATAA	145740
GTTCAATCCC	AACAACCAAA	TCTTTAAATC	CAAGGTTTGT	CCCCTCAATC	TTATCTTTAA	145800
AAAAATTACC	AATATAATCC	CGTTCGGGGC	TTTTTTGCGG	AAATCCTAAA	TATAAAAGAC	145860
GATAAAGAGC	ATTCTTTAAA	AAAGAATAAG	CATTAAGTAT	GGAAGGGGTT	TCTGCTATTG	145920
TAACAATTAC	ACCGCTGTAA	GATGCCAAAT	AAAAATCTAT	TGTATTATAA	GAAGTTCCAG	145980
ATCCCAAATy	TAAAAAATA	AAATCAGCAA	TAAGATCTTT	TTGAATGGAT	TCTATAATCT	146040
TTTTCTTAAT	AGAAAAAGGA	AGATTAGCTG	TTCCCGTATA	AAGAGCATCA	CCTGGAATAA	146100

GATAAAGCTT	ATCATAAGAT	GTTTTACATA	CTAAATCTGA	AAAACTTTTA	CTCTTTTTAT	146160
TAATAAAAGA	ACCAATGCCC	ACACCCTTAT	TTTTAACCCC	CAAACACGTA	TGTAGATTAG	146220
AGCCACCAAG	ATCAAGGTCA	ACAAGTATTA	CAGTTTTACC	CAAAC TAGAA	AGCTTATAAC	146280
CAACATTTGC	AACAAAAGAT	GTTTTTCCAA	CACCGCCTTT	GCCACTTGCC	ACAGGAATAA	146340
TTTTAGTCAT	TCTTAAATCC	TAATTATCCT	TACGATCTTT	TTGAAAAATT	TTCATAAAAAT	146400
TGAAAATCCC	TAAAAATCTA	GATTTTTTCT	CAGCATCTTT	ATTTAAATTT	TCATCCTCTT	146460
TAGAGCCTGA	AATTAAATCT	TTAATTAAAT	CTTTATTATT	TGTAAAATTT	TCAATGCTAT	146520
CTGGTTTACC	ACAAATTACA	ATTTTATCAT	CTTTTAAAAA	AAAATAATCG	CCATCAACAA	146580
ATTCATACCT	AGAATTACTT	AAATTTCTAA	CAGCAATAAC	TGTAATCCCA	CATTCTCTTC	146640
TAAGATCGGC	TTCAAAAAGA	GTTTTACCAA	CATATTCCTT	GGGAATAACA	GTTTCAGCAA	146700
CAATAATATC	ATACCCAATA	ATATTATAAG	TTGAAAGATT	TGGAGATACT	AATAATGGAG	146760
TTAATCTTCT	TGCAGCATCT	TTACTTGGAA	ATATAATTTT	TGTTGCCCCA	AGAGTTTTTA	146820
AGATTTTCAGC	ATCATCTCTA	TTTTCTGTCT	TAACGCATAT	TTCTTTCAAA	CCTAAAAGAT	146880
TACAATAGTG	AGTAACAAGA	GCACTTTTGC	CAAGATCATC	ATCAAAATCA	ATAACAACAG	146940
CGTCTGTATC	TACTGGAATT	ATTCTTTTCA	AAGCATTTTT	AGTGAATTGC	TCAACAACAA	147000
AGCTTTCTGT	AGATATCACA	TCATATTCTT	CAATAAGCTC	TTTAGATGTA	TCTATAATAA	147060
TAATTTGACA	ATCAAGCCTG	CTTAAATCTT	CAAGTAAGTG	AATGCCTAAA	TTACTAAGTC	147120
CAATAATAAC	AAATGTTTTC	ATATGCTTCA	ACCAACCAAA	ATATCTTGCC	TTGGCCTTGT	147180
AAATTCTTCA	AAACGCGACT	TTCTTGAAAC	AAAAACAGCC	ATTGAAAAAA	GCCCTATTCTG	147240
TCCTGCAAAC	ATAGTAAAAA	TTATAATGAC	TTTCCCCCAA	AATGACAAAT	CCTGAGTTAC	147300
TCCAAC TGAA	AGACCAACCG	TTCCAAAAGC	AGAAAATACT	TCATAACCTA	AATCAATAAC	147360
CTTCCAATTG	CCAGATCCTC	CCTCAAAAAA	AAGAAGCATG	AAAAAAGAAA	AACTTAAAAT	147420
AAAAATAGCT	CTTGCAAAAAA	ATAAAAGTGC	AAATCTTATA	CTATCTATTG	AAACCTTGTA	147480
AGAACCAATA	ATATATCCAT	TGCCGTTTTG	ATTTTTAACA	ACAGCCAATA	CAATTAAAAA	147540
AAATGTTGTA	ATCTTAATCC	CTCCTGCAGT	TGATCCGGGT	GCACCACCAA	TAAACATGAA	147600
TGGTAGAGAA	ATTATTTGAG	TTCTTCCGCT	TATTAAAGAA	TTATCAAGAT	AATTAAAACC	147660
AGCTGTTCTG	GTAATAATCG	AATAAAAAAT	TGAATTAAAT	ATTAAAGTGC	TCATTGAATA	147720
ACCAGCTTTT	AATTTATGCA	TCTCTGTAAA	AAAAAATAAA	ATTGCACCAA	TTATAATTAA	147780
AAAGAAGCTT	AAAGAAAAAA	CTATCTTGGC	ATGAAGCGAT	AGTTTTTTTTT	TGTTTTTAAT	147840

AGTGTTATTT	ACATCTCTAT	AGACCATAAA	CCCAAGCCCA	CCACAAATTA	TTAAAATAGA	147900
GACCACAAC	ATAGCTTCAG	GAACATCTCG	CCATGCATAA	ATACTCTCAG	AATGCATGGA	147960
AAAACCTGCA	TTGCAAAAAG	CAGAAATTGT	CGTAAACAAA	GCCTCTAAGA	ATGAAATATT	148020
CACTCCCCTA	AGTTTAAAAC	AAATAAGTAT	TAATATTAAA	CCTATCATTT	CAATTGAAAA	148080
AGTTATAAAC	AATATGCTTT	TTAAAATTCT	AATAGGATTA	TATTCTATAT	TTGAAAGGGA	148140
ATACTGCTTT	ATTATTCCTG	CATCTGTAA	ATTCATTTTC	TTTTTAGGTA	TAAGCAAATA	148200
AAAAGTAGTA	ATACTTATAA	ATCCAAGTCC	CCCAAGCTGG	ATTAGCAACA	TTATCAAAAT	148260
AAATCCAAAA	GTAGAAAAGC	CTTCCATTTT	AACCGTTGTA	AGGCCCGTAA	TACTTACAGC	148320
AGAAACAGCA	GTAAAAAGAG	CATCAATGTA	TGCTAATTTG	CCATCACCTT	CCCAGGAAAT	148380
AGGCAACATC	AACAAAAGAG	AGCCTATAAA	CATAATTAAA	ACAAAATAAC	TAAAAAGTAA	148440
AAACCTGTCG	CTAAATTCAA	ATTTCAACAT	ATCATACAAA	AAGTTGTTTA	AATTATTAAA	148500
AATTTATCTT	ATATAGCATA	ATATTTTAAC	ATTGAAATAT	TATCATAATT	ACATTATTTT	148560
TAATATATGT	TTGAAATAGA	ATCAAAAGCA	TTTATTCTTA	CAAAAGAGTT	AAAAAGAATT	148620
ATCAAGCTAG	CAAATAAAAA	ATTTAAGTTT	ATTAAAGAAG	AAATAAAAAAC	TGACATYTAT	148680
TACTCAAACC	AAAAAAAAAAT	TATAAGAATA	AGAAAATTAA	ATACTCTAGA	AAAAATTGTC	148740
ACATTCAAAA	AAAAATATT	AGACAACAAC	AATACTGTAG	AAATTAATAA	AGAGATAGAA	148800
TTCAAAATAG	ATAGTATTAA	TAATTTTTTTA	ACCCTTATAA	AAGAGCTTAA	ATTTAAAAAG	148860
CTATACAAAA	AGATAAAAAA	AAGTTTAATT	TATCAAACCTA	ACAATTTAAA	TGTAGAGATA	148920
AACGAAATAA	AAAATCTTGG	GTTTTTTTTTA	GAAATAGAAA	AAATAATTAA	CAATCAAAAT	148980
GATATAGACT	TGGCAAAAAA	AGAAATTGAC	AACATAATCA	ACCAATTTGG	ATTAAAAGAA	149040
AACATTGAAA	CTAGACCTTA	CTCTGAATTA	CTTTCATTGG	CAAATCAAAG	TAAAAAATAA	149100
TTTATTGAAA	TTAGAGCTTA	AAGTAGAGAT	TACAAGCCCT	TGATTGCCAT	AAATTCCAAT	149160
CTGAGGGCTT	TTAACATTAC	TCTTAAAATT	CTCAAGCTTA	TTTAAAAAAT	ACCAATTTTTT	149220
ATTCTTAAAA	TAAATTAATC	TCACATTATT	ATTGTCCTCA	AAAGCTAAAA	ACAAATTATT	149280
TTTATAAAGC	CCAATGTCAG	CACTTAAACC	TTCCATTTCA	ACATTAGGAC	TTATATTAAT	149340
CCATCTACTA	CTTTTCAAAG	GACAAATGTT	TACAATAGGT	CTATTTTCAG	AAACAAAAC	149400
CATAATTATT	TGATTAAAAT	TAGAATCAAA	AAAGCCTTTA	ATAAAATTGG	CCATATAAAC	149460
AGAAGGAATA	TTTGCATTTA	CCCAAGCATT	TTCATTGTTT	ACAATAAATT	CAGATTTAAT	149520
CTCATTATTT	GACTTATAAT	TATAAAAAAT	GCCCCAAAAA	GGTTCAGATA	TTAAACCAAT	149580
GTTTGATGAA	TTAACATTAG	AATCACCTTT	ACTTAAATAA	GCATGTATTA	CATCGGTCCA	149640

AATACTTCCG	TAACCCATAT	TCGAGATTAA	ATTAATTTTA	TATTCACCCC	TAATTTCCCT	149700
TAAATATGCT	AAATACAACC	TATCTTTTAA	ATCAATGCTA	ATATTTAATA	AAGATCCAAA	149760
ATTTTCTATG	TGACCAGGAC	TAATATCAAT	CCATTTTCTA	CTATTAAATT	TTTTAACTAT	149820
AAGCTCGCTG	GCAAAATCAG	CCCCTGATTT	CGTAACAAAA	GCAATATATA	AATTTCCTTT	149880
AGAATTAATT	GAAAAATCAA	AATTAECTAT	ATTAGTAATA	TTTCTATTAA	CAGATGAATC	149940
AAGATTAAAC	CAACCAACAT	CCTCAATAAA	TTCAGCAACT	TTAATATCAT	CGCTATTTTC	150000
TAGCTGATAA	GCAATATAAA	TATTGCTTTT	ATAAATCCTT	AATACATATT	TTTTAAGCTT	150060
GGCAGTTAAA	TTTAAAACAG	GCAAACTTTT	TAAAGTAAAA	AATAAATCTT	CTTTTTCAAT	150120
CTTAGATGTT	TTAGCTTTCA	GGGAATCGCT	ACTTAAAATT	GAAAAATCTA	AATCTGTAAG	150180
CGAAAACTTT	ATGTTTTTCAT	TTTTAGTACC	AACATATATT	ATTGCATAAA	GAGAATTTTT	150240
ATCTAATCTT	ATTTTAAAAT	CTCTTCTTTT	TACTTTATCG	GTTATATATT	TTTTATTAGA	150300
AATGTCATAA	ATTTTAAAAA	CAAAATCGGA	ATTTGAACTC	TTATCTAGGG	TTAAAATATA	150360
ATCGGAAGAT	TTGCTAACTT	TTAAGTAAAC	ACTTCCTTTC	CCATTTTTGC	TTAAAATACT	150420
TAAAGGACTA	ATTTCTGTTA	ATATTTGATT	TGCTTGAGCT	TGAACAAAAG	AAAATTTTGT	150480
AAATAAAAAAT	AGCAAAATGA	ATGTCTTATT	TATTTTCATA	TTTTTTTACA	TTCAAAAATA	150540
TTAACACATA	TTCTAAAAAT	GATAAAATTG	CAAAAAAAGC	AGCACAAACA	TATGTCATTT	150600
GAACAATAAA	TAAAAATTTA	AATTTAAAAG	TTAAAATGTA	ACTAATAAAA	TTTTGAACAG	150660
ACTCTGTAAA	GTTGAGTTGA	TTTAAAGTAT	AAAATAAAAAG	GCTTGCAAAA	GTGCAAACAG	150720
CATAAAGAAG	TGACTTTAAT	TTCCCCAAAA	AATTTGCTTG	TTGAACTACA	TTAAACTGAA	150780
TAATTAATTT	TCTAACAAAC	CCAATAGAAA	TTTCACGATA	AATAAATATT	ACAAAAAAT	150840
AATAGGGGGT	TATACCTTTG	TAAAAGAAAA	AAACAAAATA	TGTTAAATGC	TGCAAAACAT	150900
CCGCATAAGG	ATCTAAAATT	TTACCTACAT	TGCTAACAAG	ACCATATTTT	CTTGCAAGAT	150960
AACCATCAAT	AAAATCAGTA	AATTCATTAA	AAATAATTAA	AAACCAAATA	ATTCCAAAAA	151020
ACAAATACGA	AAAAAATACA	TTTTCCAAAA	AAAATAAAAT	TAATATGATA	AAGGAAAGTG	151080
CAATTCTAAC	TAATGTTATT	TTATTAGGGG	TAATGACCTT	GATTAAATTA	TTCAATTTAT	151140
CAAATCTCCT	TATCTCTTAT	TTTAAATAAA	ATAAATTTAA	GAGCTTCATC	AAGTTTCATT	151200
CCATTTATTT	GCTCATTTGT	TCTTGTTCTA	ATAGATATTC	TCTCTTCTGT	TGCTTCTCTC	151260
TCACCAATTA	TAAACATATA	AGGTATTTTT	TTAGCCTGAT	ATTCTCTAAT	TTTAGCATTC	151320
ATTCTTGAGG	AACTATTATC	AAGCTTTATT	CTAATCCCCT	CATTTTTTAA	TTTATTAAAA	151380

ACCTTAATAG	CATAATCTTC	CAATATTG	TTAACAGGAA	TGATTACTAC	GAACAGGA	151440
GATAACCATA	AAGGAAAAGC	ACCACCATAG	TGCTCTACAA	GAATTCCAAA	AAATCTTTCA	151500
ATAGATCCCA	ACAAAAGCTCT	ATGAATCATA	AATGGTCTTT	TTTCTTTACC	ATCCTCAGCG	151560
GTATAAGTCA	TATTAAATCT	CTCAGGGAGA	TTAAAATCAA	ATTGAATTGT	ACTCATCTGC	151620
CACTCTCTCT	CAAGCGAATC	AACTATCTTA	AGATCAATTT	TAGGCCCATATA	AAAAGCACCT	151680
CCACCCTTAT	CAATTTTCATA	AGGAAC TTCA	AAATCGCTTA	AAGTCTCTTC	AAGAACTTTT	151740
AAAGACATTT	CCCAATCAGA	ATCATTGCCA	ACAGATTTGT	CAGGCTTTGT	AGAAAGATAT	151800
GCCTTTGGGT	TGCTAAAGCC	AAATTTACTC	CACATATAAA	TAGCAAACCT	AAGAACTTCT	151860
TTAATCTCAT	CTAAAACCTG	AGAATGGGTG	CATATAATAT	GAGCATCATC	CTGAGTAAAC	151920
CCTCTGGCTC	TCATCATACC	ATGCAAAGCA	CCTATCTTTT	CATAACGATA	CACAGTGCCA	151980
AGTTCGGCCC	ATCTAAATGG	CAAATCTCTA	TAAGAATGCT	TACCTGTATT	GTAAATTGCA	152040
ATATGAAAAG	GACAATTCAT	GGGTTTAAGA	TAATAATCAC	TTTTATCCAT	TTCTATTTTT	152100
TCAAACATGC	TATCCTTATA	AAAGTCTAAA	TGACCAGAAG	TTTGCCAAAG	CCAAGATTTG	152160
CCAATATGAG	GAGTAAAAAG	AATATCATAC	CCATTTTTTG	AGTGCTCTTC	TCTCCAAAAA	152220
TCTTCTATTA	AAGCTCTTAT	TTTGGCACCA	TTGGGATGAA	AAAAACAAG	TCCTGGTCCA	152280
ATCTCTTCAT	GTATAGAAAA	TAAATCAAGC	TCTTTTCCAA	GCTTTCTATG	ATCTCTTTTT	152340
TTTATTTTCT	CTCTCAAATT	AAGATAAGAT	CTCAGTTCTT	TTTCATTATT	CCATAAAGTT	152400
CCATAAATTC	TGGTAAGCAT	TGGGTTTTTT	TCCTGCCCC	GCCAATAAGC	CCCAGCAATA	152460
CTAGTAAGCT	TAAATGCCTT	TGGATCAATT	TTATTCATAT	TCTCAACATG	AGGACCTCTA	152520
CAAAGATCAA	CAAAATTGTG	ACTCTTGTA	ATAGAACTT	CATTTTGTA	ATCAAAATTT	152580
TTAATCAAAT	CAATCTTATA	AGGTTTCATCT	TTAAAAATTT	CAAGAGCCTG	TTCTACGCTT	152640
ATTATCTCTT	TTTCAAAAGA	ACTTCCGGTC	TTTAAAATTT	CTCTCATTCT	ATTTTCTATG	152700
TCTAAAAGAG	AATCTTCTGT	AATCTGCTTT	TTAAATTCAA	AATCATAATA	AAAACCATCT	152760
TTAATAGGAG	GACCTATTGC	AATCTTGCTA	TTTGGAATA	AATCAAGAAC	AGCTTCTGCC	152820
ATAACATGAG	CTATTGAGTG	TCTTTTTTTG	TAAAGAATAT	CTTCTTTATC	TAAATCTTTG	152880
CTCACAACAA	TACCTTTTGC	CTTTCGCTTT	TTATTAAAAA	ATTAAAATTC	ACACTCATCA	152940
CTTTTACGTA	AAAATACGCA	CCTCAAATAT	TTATAATTAC	TAAATTAAAA	TATACAAAAA	153000
AATTTTCTAA	AAAAATAGAG	ATAAGAAAAC	AAAAACCTGA	AAATAAATTT	TCAATCCATA	153060
GCAACTATTG	ATTCAATATT	AAAATAAAAA	GACATTGCTA	AAAAAATGT	AATAGTAGAA	153120
GAACCTCCAT	AAGAGAGAAA	AGGAAAGGGA	ATCCCGGTAA	TAGGAAGAAC	TCCTAAAGAC	153180

ATTCCAACAT	TAAAAGAAGT	ATGAAAAAAT	AAAAGTCCCA	AAATTCCAGA	TATTACTAAG	153240
GCCATATATC	TATCTTGACT	TTTATTCATT	ATTATCAAAA	ATTTAAAAAA	AAGGAAAAAA	153300
AATAATATTA	AAATAGTGCT	AACACCCAAA	AACCCAAACT	CTTCGGCAAG	AATAGAAAAA	153360
ATAAAATCTG	TGCTTTGAGA	TGGCACATAA	TTAGCGTGGG	TATAAGGTCC	CTTTAAAAAT	153420
CCTTTGCCCA	AAAGACCGCC	AGAACCAATT	GCTATTTTAA	CCTGATTTAA	ATTCCAACCA	153480
GCACCTTAG	CATCAATAGC	CGGATCTAAG	AATACCAAAA	ACCGTTTAAT	CTGATAAGTC	153540
TTCATTAACT	TTGAAAGAAC	CTTTGAAAAC	ACTATTGAAA	CTAATAAAAT	AGAACTTGCA	153600
AAAAATACAT	AAAAATAAAT	TATTTTAATA	CTCAAACCAT	ATTTAGAAAT	GAAAAATCCT	153660
AAAACAGAAA	TCAAAAGAAT	TAAAAGCAGC	ACTCCCATTA	TTACTCTAAA	ATAAAAAGGA	153720
TTTGAGAAAA	TAAGATAAAA	TACATTACCC	ATATTCACCT	TATATTCATA	CCAAACCGGT	153780
AAAATTGCAA	AAACAAAAGA	AAAAAACCTT	ATCAACGCAA	ATGCTAAAC	ATAGTGCAAA	153840
TCTATTCCTG	CAAAAAAGA	AATAAATATA	AAAATGGTTA	AATATACTAT	TGCTGTACCA	153900
AAATCAGGTT	GCAATAATAT	AAGAATTACC	GATGGAAAAA	TTAATAAAAA	TGCAGTAATA	153960
AAGGTAAAAA	ATTCATTATA	ACCCTTTTTT	TCAGTGTAAG	ATTTTGAAAG	GGTTAAAATA	154020
ATAACAACCT	TACCAAATTC	AGAAGGCTGT	CCTCCAAGTT	TCCATATGCC	AATCCAAGAT	154080
CTTGCTCCAT	TTACTGTCAT	TCCAAAAAAT	GCAGTAAAAA	TTAAAGCCAA	TATTAATAAA	154140
AAATATAAAG	GATATACCAT	GCTATAAACA	AATTTTAAAT	CATATTTGCC	CACTATAAAA	154200
ATTAGAAAAA	ATCCAATAAT	TACCCAAAAG	GTTTGTTTTT	TATATTCATT	CTTGGTTAAA	154260
GATCCACTAA	TATTATAATC	GCTAGAATAA	ATCAACAATA	TACCAACAAA	AGAAACTATA	154320
AGTAAGCTTA	TCAAAGCCAA	ATAATCATAA	TTTTTTCTAA	AAACCATTAA	TCTACCTAAT	154380
ATACCACGGC	CTATAACCTT	TAAGAATATC	TTCATAACTT	TGATTTGCAA	AAATGCCTTG	154440
CATTATTAAA	TCTGTAGATT	TTGCAGGCCA	CCAATCCACA	TTACTTTTTG	CCTCAACCAA	154500
ACTAAAAACA	ATAATTTGAT	TATCAGCTGA	ACCGTTATAA	GGGGCAAGTC	CAATAAAAGA	154560
ACTATTTTCA	AAACCATCTA	TTCCAGTTTG	ACCAGTACCT	GTTTTTCCTC	CAACCTCAAC	154620
AGCTTTGGTA	AGAACTGCAT	ATCTTGCTGT	ACCATAAGTT	ATAACACTTC	TCATATATTT	154680
TTTCAGAAGT	TTAAATGTGT	TTTTACTAAT	AAGATTTGTC	TTTCTTAATA	TTTCTGGTTT	154740
ATTTTCAAGA	ACAACCTTAT	TAGTACCACC	TTTTAAATTT	TTATTTACAA	TTCTAGGTTT	154800
ATATACAACA	CCTTCATTTG	CAATCATAGC	AACCATATTA	ACAATCTGCA	TAGGAGTAGC	154860
ATTTAAAAAT	CCTTGACCTA	TTGAAAAATT	TACAGTATCT	CCTCCTACCC	AAGGCTGATT	154920

AAAAGTTTTT	TCTTTCCACT	CAGGACTAGG	AAGAAGGCCA	GCTACTTCAT	TTGGCAAATC	154980
AATTCCTGTT	TTTTCTCCAA	ACCCAAATTC	TTTTGCATAT	TTTCTAATTC	TATCAACTCC	155040
AAGATACTTA	AGCCCAAGTG	TATAAAAATA	AACATTAGAA	GAATGTGCAA	TCGCCTCTTC	155100
TAAATTAACA	TACCCATGAC	CTCCGGGCTT	CCAGCAATGA	AAAATTCTAT	TTCCAACTTT	155160
AAAATATCCA	GGACAATAAA	TTTTACGATC	TTTGTCTATA	ACTCTTTCTT	CAAGAATGGC	155220
AGCAGCAACA	ACTAATTTAA	AAATAGACGC	AGGCGGGTAA	ACAGATTGAA	TTGCTTTATT	155280
TAAAAAAGAG	TAATCTTCCT	TATTATCTTT	ATTGTAAACA	TCTTTCATAG	AATAATAAGG	155340
ATAATTGTGA	AGAGCAAGAA	CAGCACCTGT	TGATGGTTTT	AATACTACAA	CAGAACCATA	155400
CCTTTTGCCT	AAAGCATTCT	TAGCAAGATC	TTGAATATCT	TTATTGATAT	TAAGCACAAC	155460
ATCATTACCG	GGCACCATAT	TTTTTATAAT	AGAACCATCG	TCTATTCTTC	TCTCCTTAGA	155520
ATCTACCTTG	TATTTTATTA	ATCCCTCTTG	CCCTCTAATG	TAATTATCAT	AACTTGTTTC	155580
AACGCCCAAC	TTTCCAATCG	TAGAAGTATT	ATCATACCCA	CTAACATTGT	AAAACGTCCT	155640
AAGTTCTCTT	TGATTTATTT	GCCCAACATA	ACCGATTGAA	TGAGAATATG	AATCGTCAAC	155700
TAAATAGTTA	CGCTTAAAAG	AATAGGTCCA	CAAAAGAGCA	GGATAATAAA	ACTTTTTTTTC	155760
AGAAATTTTG	AAAAGCATCT	TTGGGGTAAG	TTCAATTATT	TCAACATCTT	TAAGATATCC	155820
ACCAGGCTCT	TGAAGTTTAG	ACAAAATAAT	TGATTTATCA	ATATCTAGAG	TGCTTGATAA	155880
AAAATCTATC	ATCTCAATTC	TAGTAGCAGC	AGGCATATTG	TAATACTGTT	GTAAGCTTAT	155940
CTTTAAGATA	AACATAGTTA	AATTATTTGC	CAAAACATTG	GAATTAGAAT	CCAAAATTTTC	156000
ACCTCTTGAG	GCATTGATTT	TTTCCAATCT	TGATAAAAAA	ACATTGGCTT	CTCTGTCATA	156060
AAACAAATGC	TTACCAATTT	GCATTTGGAA	TAAAATCGCC	AAATAAAGCA	CCATAATTAC	156120
TATTAAAAAA	AATATGCCGA	ACTTGTATCT	AAAATTTGTT	ATAACACCCA	CTAATAATCC	156180
TCTTTAAAAG	AATAAAAATT	TCTAGTAAAA	TAATTTTGAA	TTGGATATAA	AAAGTTAATA	156240
GACATTATAT	TTACAAAAAG	ATCAAGGTTG	AAAATTGAAT	AATTAAAAGA	TTTTAAGTCT	156300
ACAAAAATCAT	AAAACACAAT	AGCTAAAAAC	CATAATATAA	TTTTTGAAAG	AATAAAAAAT	156360
ATTGTCATGC	TAAGCATATT	TTTGGGCATG	AATAATTTTA	TTTTATTGTT	AAAATAAAAT	156420
ATTATCGTAT	ACCCAAAAAC	AAAAAATCCA	AGTGGTAATC	CTGTAAAATA	ATCCATAAGA	156480
AGACCATATA	AAATGCTAGA	TAATAATCCC	ACATTAAAAA	TAAAATTCAA	AGAATTAAAA	156540
ACTAGAAAAA	TTAAAAAAAT	ATCTATTGAA	AAATAAAAAAT	AAGTTGCAAA	ATAGTGTTGA	156600
AAAATTTTGC	CTAAAAATGC	GCTGGAAATA	AAATATGTAA	AAAATGTTGC	CATTATTCAC	156660
CAATCTCTTT	GTTGTTTTTA	ACAAGAAAAA	CATACTCAAG	CTTATCTAAA	ACTATAGCTG	156720

GCTCTACTTC	TATTTTAA	AGAGAATTAT	AATCAAGAAT	ATGAAAATTT	GTAATCTTTC	156780
CAATATAAAT	ACCAACTGGA	TATTCACTAA	ATCCAGCAGT	AACAATAGAA	TCCCCTATTT	156840
TTAAATCTTT	TTCAGCAAGT	CTATTAACGT	AATTCATTTT	AAGTTTTTTA	CCATAACCAT	156900
TGCCTTCTAT	AAGGCCTATA	AACCTACTAC	TTTGAATCCT	TGCGGACACA	AAATTTTCAT	156960
AATTAGTTAA	AGGCAAAATT	TTAGCAGTAT	TAGAATAAAC	CTTTACAAC	TTGCCTACAA	157020
GGCCACTAAA	TCCATCCTGA	TATGCAACTG	CTATCATATC	TTTTTCTATC	CCATCATTGA	157080
ATCCTTTATT	AATAGCCATT	AAAGTCGATA	TGTTTGAATA	GTTTAGATAT	ATAATCTCTG	157140
CCGAAATAAA	ATCGCTAGAG	CTTGACGAAT	AAAAATTTAA	TTGCTCTTTA	AGACGAACAT	157200
TCTCTTGCTT	TAGTGACTGT	ATATTCCTGAG	TGACTATTTT	AAGCTGTTGT	ATCCTTTTTT	157260
TATAAAATTC	TATCTTGTCC	TTGTAATTTT	TGTATTCATT	TACAGTTTTA	AAAACATTGG	157320
AAATAAAACT	AAAAACCCCA	TGCATTCTGC	TTTGAATATA	AGAATTAAGA	GTAACAAACA	157380
AAAAATTATC	TGATCTTCTC	TTTTGAATGC	TGCTTGAATC	ATGAATCATA	AAAACAAGAG	157440
AAACTATCAA	TACCAAAAGT	ACTTTGATAA	AATTCTTGAA	TTTGACAAGA	AAATTCATAA	157500
CTTATTCATT	GATAAACTG	TAAATATTTT	TACTAATATC	TATTCTATTG	GCATAATCAT	157560
AAAATAACCC	GGCACCAACA	GCTACCGAGA	GAAGCGGATT	GTCTGCAACA	TAAACAGGAA	157620
CTCCAGTCTC	TTTTGAAAGA	AGTCTATTTA	AACCCTTAAG	AAGAGCCCCT	CCCCCTGTCA	157680
AAATAATGCC	ACGCTCAACA	ATGTCTGTAG	CAAGCTCTGG	GGGAGTTGCA	CCAAGAGTGC	157740
GCTTAACTTC	ATCCACAACA	ACATTTATAG	GTTCTTGCAA	AGACTCTCTT	ACTTCCATAG	157800
AATCAACAAG	TTGCTTTCTA	GGAAGACCAG	TTACAGCATC	TGTACCCTTA	ATGTCTATTT	157860
TTTCTACCTT	TAAATTTTGA	ATATCGGGAT	ATACATTTCC	TATCTTAATT	TTCAATTTTT	157920
CTGCTGTCTG	TTGACCAATT	ATAATATTAT	GAGAATTTCT	CATATACTTT	ATTATGCTCT	157980
CATCAAATTC	GTCACCACCA	GTCTTAATTG	CTCTACTTAC	AACCATGCCG	CCAAGAGAAA	158040
TAACAGATAT	TTCTGTAGTT	CCACCCCCAA	TATCACACAC	CATATGACCT	GTAGGTTCAA	158100
AAATAGGAAT	ATCAGATCCA	ATAGCAGCTG	CAAGAGATTG	TTCTATTACT	TTAACTTCTC	158160
TTGCACCGGC	ATTCATTGCG	CTCTCTTTTA	CAGCTCTTCG	CTCAACCTCT	GTAATACAAG	158220
TTGGAACACC	TATTACCATT	CTCGGCTTAA	AAAATAATTT	TTTACGAGAA	AAAATTTGAT	158280
TAATAAAATA	TTTGATCATC	TTCTCTGTAT	TCTCAATGTC	AGCAATAACT	CCATCTCTAA	158340
GTGGGCGTAC	GGCTTTAATA	TTTTCTGGAG	TTTTCCAAAG	CATTTTTTTA	GCATTTCTAC	158400
CAACCGCAAC	AACTTTATTA	CCTTTGGTTA	TATCTATTGC	AACAACAGAA	GGCTCGCTCA	158460

TAACCACGCC	ATAATCTTTA	ATAAAACCA	ATGTATTACA	TGTTCCAAGA	TCAATGCCAA	158520
TATCTATCAA	AAAAGACTTA	AACAAATTCA	AAACAACCTC	CCTAAAAGTC	TTCCAAAGTA	158580
ATTCCAAGCC	TCTCTCTTGC	TGCTTCTAAA	TAAGGATTAA	GATTGAGAGC	TTCTCTCCAG	158640
TATTTTCTAG	CTTTAGGATA	ATCTTTATCT	TTCTTTCTAT	ATATATCTCC	TATTTTAACA	158700
TAAACGCTAG	AATTTGAACT	ATTAATTTCT	AGAACTTTAT	TATAGTAATT	AAAAGCACTG	158760
TCATAATCGC	CTTTATCTAA	TAGAATGTCT	CCATAAAGCA	AATATACCTT	TTGAATTAAA	158820
TTTTCATCCG	TTTTCTCACC	CTTTGATAAT	TTTCTTTTCT	CTTCTTCTAT	TATTTTTTTT	158880
ATATACTCAA	TGCTTTTGTT	TATATCATTC	AAC TTATAAT	TAACATAAGC	CAAAC TCCAA	158940
AGCACTAAAT	CAGATTTATT	TCATTAAAA	GCTTTTTTCAA	AAAAC TTCAA	ACTAGATTTG	159000
TAATCTCTTA	AAAGCTGATA	CGAATATCCC	AAATATTCAA	AAATATCTTC	TCTAATGTTC	159060
ATGAAATCAA	AATTATCGGC	ATTTAAGGCC	TTCTTTAAAA	ATTTTACAGC	AAGCTCGCTA	159120
TAAAAC TCTC	CTTTATGAGA	ATATGCCTTT	CCCAATATGT	AATACAAAGG	GCTTATGGAG	159180
ACTCCATCAT	TTATAGAAAT	TAAAAATCTT	AGTCTTTCTA	TGGATTTATC	TAAAAACTCT	159240
CCTTTTAAAT	ACCCTTCATT	TACTATTAAA	GAATAATAAA	AATATGAAAA	TCCTAAAAGT	159300
AAATTCAAAT	TAAAATCAAA	TCTATGATTT	TTAATGTCAT	TCTCAGCATA	ATCTATTATT	159360
TCTTTATATT	CTTTTTTATC	CCAGAGTAAA	AGCAAATCAA	CTTCTGTTGG	ACCTGCCTTT	159420
AAATAAGAAC	TAGAAAAAGA	TTTAAAATAT	GATAAAATGT	AATAAATTAA	AAAAATGAAA	159480
ATGAAAATCA	TAAATGAGTA	AAAAATATAT	CTTAAATATC	TTATTTCCAT	TTAAAACTCT	159540
CTTTTAGGTC	AATGAAATTA	AGCTGACACC	CATAAGCCGA	GTTCTGTACT	ATGCCATCAT	159600
CTCTCTTATT	TTTTTATCGC	TAAAAAAATC	GTGCGATCTA	CCCGTAAGCA	TGTCCTCAAG	159660
AACAAAGGGT	GCTTACATAC	TTGATCTTGC	TCCTAATGAG	GTTTATCTTG	CCTGTATTTA	159720
TTGCTAAATA	AGCGGTGAGC	TCTTACCTCA	CCTTTTCACC	CTTACCTTTT	ACGGCGGTAA	159780
TTTTCTGCGA	CACTTTCTTA	GGTTTAAAAC	CCCTAGGCAT	TACCTAGCAT	TATGTTCTTA	159840
TTGGAGCTCG	GACTTTCCTC	TTAAGCTTTA	ATTATAAAAC	TAAGCGATGG	CTGACTGCCA	159900
GCTTAAATAA	AAAGTATCAA	ATTAATAAAT	TTTATTCAAC	AAGATCTGCT	AAACTACCAG	159960
GAATTATTTT	ATCGCTTATC	TGAACTTTAT	CCTGATAATA	AAGTATTCTG	CTGCAATAAG	160020
GACAAAATTT	AATATCGTTG	GGCTCACGTC	TTACTTTTATT	TGCAAATTCA	ATAGGAAGTA	160080
TCATATGACA	ACCTTTGCAA	ACATTGTTAA	CCAAAGGCAC	AACTCCATTT	GATTTATTTT	160140
TTATTATTCT	TTGAAATTTA	AATAAAAAAT	CTTCATTCAT	TTTAGAAGCA	CAATTTAACT	160200
CTTCACTCTC	TATTTCTAAA	AGTTTCTTTT	CAATTTCTAA	AAGCTCCAGC	TCAAAACTAC	160260

TGCTTTTCAGC	TCTAAAACAT	TCCTCTTCTT	TGATGTGCTT	CTCGTTGACA	TCTAATATTT	160320
CCTTTTCTAT	TTTAGTTTTA	AGCCCATTTAA	CATGTGTCAT	CTTTTTTCTA	ATTGTAACCTT	160380
CATCGTCAAT	AATAACCTGA	AGTTCTTTTT	CAAGAGCCTC	ATATTCTCTT	TGCGTTTTTAA	160440
TGCTATCAAT	TTTTTCTTCA	GCCTTGCTCT	TTCTTGAATT	AATATCTTGA	ATATCTAACT	160500
TTAAAGCAGA	GTCTTCTTTT	TGATACTCCT	TAAACTTTTG	TTGCAAATCA	ACAAGAACTT	160560
TCGACAATTC	TTCAATCTGA	TTTTTTTTTCG	CCTCCAAATA	CTTGGAATA	CTTTTTTCGCC	160620
TTTCTTCAAG	CTCAAACTTA	GATTTATATA	TAACCTCAAG	TTTTTTTAAT	GTATCAATAT	160680
TGTTTTCCAT	CAATCCTCCT	GTTCAATTTA	AATCTTCAAG	ATAATCTTTT	AATTTTTGAG	160740
TTTTTTTGGG	ATTTTTTAAGC	CGCCTTAATG	CTTTAGATTG	AATTTGCCTA	ATTCTTTCTC	160800
TTGTAACATT	AAAATGAAGT	CCAACCTCTT	CAAGAGTTAA	AGAATAGCCA	TCTTCAAGTC	160860
CAAATCTCAT	TTTTACAAC	TCTTGTTCTC	TTTCAGGAAG	AGTTCCAAGA	ATTGCTCTTA	160920
TTTGATCTTG	CAAAACTACA	AAAGATGTGT	GATTTGCAGG	ATTTTTTATT	GCCTTATCCT	160980
CAATAAAATC	GCTAAGAACA	GAATCTTCCT	CTTCTCCAAT	TGGTGTTTCA	AGAGAAACAG	161040
GTTCTCTTGA	AACACTCTTT	ACAGTTTTTAA	CCTTTTTTAAG	TTCCCATCCA	AGCCTGTCTG	161100
AAAGCTCTTC	ATCTGTGGGA	TCTTTGCCTA	AACTTGAAT	TAAATATCTA	GTTTCTCTAT	161160
TAAGCCTATT	TATTTGCTCA	ATCATGTGCA	CAGGAACTCT	AATTGTGCGA	GCTTGATCAG	161220
AAATAGATCT	TGTTATGGCT	TGTCTAATCC	ACCAAGTAGC	ATAGGTTGAA	AACTTAAAAC	161280
CTCTCTTATA	TTCGAACTTT	TCAACAGCCT	TAATCAATCC	AATATTGCCT	TCTTGAACAA	161340
GATCAAAAAA	ATGAAGACCT	CTATTTGCAT	ATTTTTTAGC	AATGCTTACA	ACAAGCCTTA	161400
AATTAGCCTT	AATCAACTGA	TCTTTAGCAT	GCTGCATCAT	TTGCTTCCCT	TTAGCAATCT	161460
CTTCTGACAT	GCTTATTATT	TTATCAGTTG	GATATTGATA	ATACATCTCA	ATTCTCTCAA	161520
GTTCTTTTTG	GGCAAGCTGA	GCCTCTGTAA	TCTGCTCTTT	AATAGCATCT	TCTTTAAGCT	161580
TGAGAGATTT	TTCTATCTCT	ATTTTTTTTT	CAGCAATAGT	CAAATCTCTT	CCAAGCACCC	161640
TCAAATCTCT	TATTTTTTCA	ATTTTCAGCC	TGCTTAGAAT	TATCTTTTGT	TGTCTTGTGA	161700
AATCTTTTAT	TTTGTTAGCA	GAGTCAATAT	AATCATCTGA	GAAAATCCTT	AATTCCTCTT	161760
GATACAAAGG	AATGTCTCTC	AAAAGCTCTT	TTAGGGCCAA	TCTTTCCTTT	TTTAAATTCT	161820
TTTCAAAAAT	ATCCCCCCCC	AGATCATACA	CCCTATGCTT	ATTATCTACA	TAACCTATTA	161880
AACGATCTTG	AATTGGCTTT	AAAGGAATTT	TGTAAAAAGA	GGCAATTCTT	TTTTTTTTTAT	161940
TATAATAATC	CGGACTACTC	TCTTTATCCT	TATCTTTTTT	TCTTTTAAAA	AACTCTTCTC	162000

TTTCCATTCT	TGAGTAAATA	ATTTCACAA	GATTATAATA	ATTTTCTATA	CAAGTCCCT	162060
CATTCTTAAG	AATATTCTCA	ATTATACTCT	CTCCAGAATC	CATTTGCTTT	GCAAGTTCAA	162120
CTTCTTGATT	TCCCGTTAAT	AAAACTCTT	TTCTATTTC	CTTTAAATAA	AGCTTGATTG	162180
GATCTTCTGA	GTGACTATCT	TTTAAAACAT	TGCCTTTAAT	ATACCCTGAA	CCTAAATCAT	162240
CCTTAACAGA	AATATCTTCT	TCATCACAA	CATCCAGCTT	AACATCAATG	TCAATATCTT	162300
CCTCATCACT	TTGAAAACCA	TCATCTAAAA	TCATAAAATT	TCTATCAGAT	TCAATCTCAA	162360
CCTCTTCCTC	TTCATCATTT	CCATCTTCAC	TGACAACCAG	ATCTAATTCC	GAAATTTTAT	162420
TAACCAACCT	TATTCCTTA	TCCTCAAGTA	CCGAACAAAT	ACAATCAAGA	ATCTCTGGTT	162480
CTAATATATC	ATCGGGAAGC	AAATTTGATA	ATTCACTAAA	ACTAAGAGAT	TTTCTATCTC	162540
CCAAATGAGT	AATAATACCC	TCTATCAGCT	TCGAATATTT	CTTTTCCAAA	TCCGACAAAA	162600
CCTAACTCCC	TGGAACATCA	TCTATGTAGA	TTTTTAAATT	TTTCTCTGCT	ATATTTAAAA	162660
ACATTAACTC	ATTTATTTGA	ATCTTAGCAT	TTACCAAAGA	GTCCCCATCA	TATCTTTTTT	162720
TGCAAAGCAA	AACACGAGAA	TCTAATTTTC	TTCTCTTGAT	TGCAAGTAAA	ATATGAATGA	162780
GCATCTCATC	ATCCACTTCA	AATTCAGAAT	TTAAAATTTT	TTCAAAAAAA	ATTCACTAAC	162840
TTTATAGGTA	TCCTTTAAAT	TTTTTTTTTAA	ATCCATTAAT	GAAAAATCTT	TATTATTTTC	162900
AAATAAATTT	TCAAAGCACA	TAAAACTTT	TCTGGCATCG	ACATTAATTA	AATCACTATC	162960
AATAATATTG	CGCCTTACTA	TGCTAAAATA	ACTAAAATTT	TTCAACAAAG	CTACTATTAG	163020
ATACCTCTCA	TAAGAATCAT	CATTATGAGC	ATACAAATTT	CTTTTATTAT	TGTCAACTAC	163080
AAATCTTTCT	TTTATTCTGT	AATAATCTTT	CAATAAAGTT	GTCACACCAA	TACCAAGTTT	163140
ATTGCTTAGC	TTGTCTAAAA	AAATTTTTTTT	CTGAGTATCT	ACTTTTGATA	AATTTATCAA	163200
ATTTAAAAAT	AAATTAATCA	TGGCATTTAA	ATCTACAGTT	TTATTTAAAT	TATATTTATT	163260
AGAATAAACA	TCCAAAAGAT	ATTCAAAAGC	ATCACATCTA	TTATTTAAAA	TTTTTTGCAA	163320
GGAGTCTACA	CCCTCACTTT	TAAGAACATC	TGCAGGATCA	GTACCAAAAT	CCATTCTGAAC	163380
AACACTAACA	TTGATATTAA	ACGGCAAACA	AATTTGATAA	GCTTTTAAAG	TTGCAGAAAG	163440
TCCAGCATCA	TCCCCATCAA	AAGAAAATAT	TATCTCATCA	GCATATCTTT	GAATTAAAGC	163500
TAAATGCTCT	TTTGAAAAAG	CAGTGCCAAG	AGTAGATACG	GCTCTCTTAA	TCCCAGATGT	163560
AAAAAAAGCA	AGAACATCTA	TATACCCTTC	TACCAATATA	ACTGATTTTG	TAGATTTAAT	163620
CTCCTCAAAA	CCCTCATAAA	ATCCATAAAG	AAGCTCCCTT	TTTTTAAAAA	CTTCAGTTTC	163680
ACCTAAATTA	ATATACTTAG	AACCTTTCCC	ATCTAAATCT	CGACCTCCAA	AACCAACAAC	163740
GTTTCCTTTA	AAGTCTTTAA	TTGGAAAAAT	TAATCTTTGA	AATAAAATAG	AACTTTGGG	163800

ATTGGTTTTTC	GAGAACAAAC	CACTTTTTTCT	AAGTACTTCA	GAAGAGTATC	CTTTTGAAAC	163860
TAAAAAATCA	TGAAGCTCTA	AACCATTTTT	AAAGTTAAAT	GGCAAATAAC	CAAGTTCAAA	163920
TAAATCAACA	ATTTCCTTAG	ATATTGCTCT	ACTCTTTAAA	ACATAATCTA	AAGCTTTTTT	163980
GTTTTTACTT	AAAAAAAATT	TAATGGTATT	AATTAACCGA	GAATTCAAAG	AGTAAATTTT	164040
TGAAACCATG	TCTTTATTTT	CATTTTTTATT	TTCACCTCCT	CGACTTATTT	TTAAATCATC	164100
ATAATGAATA	CCGGATTTTT	CGCATAAAAT	CTTAAGAGCA	TCATTGTAAT	TGATTTTTTC	164160
CATATCCATT	AAAAATCCAA	TAACATCTCC	ACCCTTTTTG	CATCCAAAAC	AATAAAAAATA	164220
TCCTTGCAAA	GGATTTACAA	AAAAAGAGGG	AGTCTTCTCA	GCATGAAAAG	GACAAAGACC	164280
TTTGTAAGCA	GATCCCGATT	TAACAAGCTT	AATATATTGC	TCCACAATAG	CTACAATATC	164340
AAATTTGCTT	TTCATTGAAG	CTACAGTTTG	TAAATACTTC	ATACTTCTTA	ATCCTTAGTA	164400
ATAAAATTTT	TAATATAATC	TTTTGCCCTT	AAAAGATGTG	AAGAATATTC	TGATGAAAAT	164460
TGATGTGTGC	CCAACCTAGA	GTCTTTTACA	ACAAAAATA	AATATTGCGT	ATTTTTTGGG	164520
AAAAAAGCTG	CTTGCAGCGA	AATAATACCA	GCATTTGAAA	TTGGAGTAGG	AGGATATCCT	164580
TTATTAATAT	ATGTATTATA	AGGAGAATCT	ATCTCTAAAT	CTGAAAAATA	AATTCTCTTA	164640
GGATGGCTTC	GTCCTAGCTC	CTCTGTAATA	ACATATTCAA	TAGTAGCACA	GGATTGTAAT	164700
GCCATACCAG	ATTTTATTCT	ATTATAAAAA	ACCGAAGACA	TTATTGGGGC	TTCACFTTTA	164760
ACCCTATATT	CACGTTCAAC	AATAGATGCT	ATTATTACCC	TATTGTAAAG	CTCCTTACTT	164820
GAATAATCGC	TAAGAACAAC	GCCATATAGAC	TTAAGCTTAT	TCAAAAAATT	ATCAACAAAC	164880
ATGCGAACTA	CATTCTTTAT	TTCTATACCC	TTATAAAATT	TATAAGTATC	TGGAAATAAA	164940
AATCCTTCAA	GAGAGTCATA	ATCAAGCCCA	AGCTCATAAA	TAAATGATTT	TTTGTTGATT	165000
AAAAAAAGAA	AATCTTGAAC	ATCATCAATA	ACAGAAAATT	CCTTAAGCTT	TAAAGCAATT	165060
CTTCTGCTAG	TATACCCTTC	GGGTATTGTA	ACATCAATAT	TTACGTTAGA	AGATCCCTTT	165120
AAAAACTCTT	TATATATTTT	AAATGTAGAA	AGATCGCCAT	TTATTAAATA	TTTCCCCCTT	165180
TTAAATTGTT	TATCACTACC	TAAAATATAT	GAAATAAAAA	CAAGAAGCAG	CTCGGATTTA	165240
ATTAATTTTT	GTTTTTTCAA	TTCTTTAGCT	ATTTTTTTTAA	CTCCCCAACC	TTTTTCAATA	165300
TTAAATTCAT	AAACTAAGCC	ATTTGCCAAA	GAAGATAAAT	TTAAAAAATA	TATAAAAAATT	165360
GACAAAATAG	ATCCCCAAAA	GAAAAAAGA	ATAAACACTT	TCCAATTTT	AATAAGCACA	165420
AGAATCTCCT	AGCTAAACAA	AAGTCTTTGC	TAATTACACT	TTAGCTATCA	ATTAAAATTT	165480
TTAATTGATA	GCTAACATCA	TATTTTATAA	CCAATAATAA	CTTTTCTAAA	TAGACAAAAA	165540

TAAAAGCACC	AAAATGAAAC	GACATAACAC	ATTTAAAAGA	TCACAATAAA	ATTATATAAAA	165600
AGGCAAACGC	TTATAATTTT	ACTAAATTGG	GGAGGGTGGG	ATTCGAACCC	ACGTAGGCCAA	165660
AGCCAACAGA	TTTACAGTCT	GCCCCGGTTA	ACCACTTCGG	TACCGCCCCA	TAAAAGCCGA	165720
CTGTCCGATT	CGAACCGACG	ACTGCGGTTT	ACAAGACCCG	CTGCTCTGGC	CAGCTGAGCT	165780
AAGTCGGCAA	ACAAAACTA	ATAGTTAGTT	TATAAAAATA	AAATTCATTA	GTCAATAGAA	165840
TAATGATAAA	TAATTAATAT	AATCTTCCTA	TTAATCTATT	TATTACCAGA	AAGAAAAATT	165900
TTAAAAACAA	GAGCAATAAA	AAAAATTATA	AAACTAAAAA	TTAAAATTAA	ATAATACTTA	165960
ATTTTTTTTAT	TTTTCAAAAA	ATAAAAAAAT	ATCAGTCTAG	TAGATTCAAG	TCCAAAAGAA	166020
ACAGAGCTTA	AAATAATAGC	AAGATCAAGA	TAAAATTTCT	GAAAATTGTA	AACAAAAATA	166080
TTGTAAGAAC	CAGCCAAAAA	CACGAACAAA	CTTAACAAAG	AATTTAAAAA	CAATACAATA	166140
ACCAAAATTC	TAATAAAATT	TGAAAAACGT	CTTAGATTTT	TTACTTTAAA	AGCTTTCATG	166200
AAATTTATAA	ATTCTTTATT	ATAAAAAGTG	CTTGATTTAT	GATTTCCAAA	AGATTGCCAA	166260
TAACATCGGC	TAATGTAATT	GTTTTATTTT	CAGTTACCAT	TGATTTGTAA	GAAACCGTTA	166320
AAGATTTTCG	AATAAAAAACA	TTATGATTTT	CCAACAAAGA	CAAGCACCTC	TTTTTAAGCT	166380
CAATAAAATA	ATCCACAAAT	CTACTTGAAA	CCTTATCAAT	AGAAACGGAT	TCAGTTAAGC	166440
ACCACTTCAA	AATCCTTTCC	TTATAAGAGT	TATCAATTTT	ACCGTCTCTT	TTATACCTTT	166500
CAAACCTCTC	ATTTTTTATA	TCAATCTCAA	AGTAAATATT	ATAAATTTCT	TTTTTCAAAG	166560
CAGTAATATT	ATTTTCAATG	TCAACATATA	TATTTCTCAA	ATCATGATTA	TTGATTAATA	166620
CGACATTATT	TATAGTTCTT	AAAACCTCCA	TCATTCTAAC	ATCATATATG	TTTTTAAAAA	166680
AATCATTAAG	AAAATATAAC	TTTTTAAAC	TATAACGCCC	AAGATTTAAT	TCATTAAAAA	166740
TCTCGCCTTT	TAAATTTAAA	ATTAAACTAG	GATCTCTAAC	AAGATTTTTT	ATTCTTTTGA	166800
TAACAAGAAC	TTTTAAAGAG	CTATCACGCT	TTTTAATAGC	CTGACCCATA	ACATTTAAAA	166860
ATAAAATGTT	CTCATAAAAA	TTTTTTACAT	CAAAAAAGGG	TACATATTTT	TTAGGAGATG	166920
CAACAGGATT	TTTGCAATAT	ACTCTAAAAA	TATCTAAATA	AGGCAAAGTC	CTCGACAAGC	166980
CAACAATCTT	AGATACAATA	GATAAAACAT	TTTTAAATAC	ATATTCTACC	CGAGATATAT	167040
GATCTTCTCT	TGAAATTATC	TCTGAATTAC	TAATATCATC	AGATATTGAA	AAATAATTCTG	167100
ATATAATGCT	TTTAACAATA	TCCTCATCTA	TCTCAATATC	TTTTAATGTA	TACAAAATGT	167160
CAAAAAAAGA	ATTCAAATAC	TTACTAACAC	TTGAAAAGCT	GGTTCCAAAT	CCAACCTCAA	167220
AGTCAACAAT	GCTAATATTC	TCAGCACTAT	CTAAAAGATC	TACATTAAAA	AATGAAAAAA	167280
AAGATTTGTA	AGGGAAAAAC	GCCAAACTAT	TAAAAACATA	ATAAACTCA	AACATATCCG	167340

AGACGATTTT	GTAAGTACTA	GATGGAATAG	AATTAATATA	TTCATTTATC	TTGACCTTAA	167400
GAGATTCTTC	TAAATCGTTT	TGGGTTTTTT	CTTTTTTTAA	AAAAAGCTCA	TATTCACCTT	167460
CATCTAAAAA	ATCTTCCAAA	TTGTTCTTAG	AGTTAAGAAC	TTTGCTTTGA	ATAATTTCTA	167520
AAATAGTTTG	CTCAACAACA	ATACCACCTT	TTTCTAATTT	CCTAAAAAAT	TCTTTCAATT	167580
CCATTGAATA	TCGGCAAACA	CTAAAAAACA	TTTCCACAAA	ACCAACATGC	AAACATTCTC	167640
TCTTAAAAATC	TATTACTGGA	TTTTTACACA	TTTCATTAAT	TTGATTTTCC	AGGTTTTTTA	167700
TAACATAAGC	TTTATATATA	TCTTCTTTGC	TTTTATCTTT	CTGAAAAAAA	TTTAAAATAG	167760
TTAACCACAT	TCTAATAAGA	AATGATTCTT	CCATTAAGCG	CCCCGAAAGA	AATTTTTTCAA	167820
GCTGAGAATC	AACATAAGGC	TCATTACTAC	TATTGGGCAA	AACAAAATCG	GAAGAATTCA	167880
AACCTAAGCT	TTTTCTTATA	TCATTCATCA	ACCTATCTTT	AGTCTCTTTT	GAAAGTTGAC	167940
CTATAGAGTT	GTATACGCTA	TTGTTATTAC	CACTCATAAA	ACCCCAATAA	AACCATTTAA	168000
AAAATAAAAGT	ACTACTTACA	ACATACTTCA	AATAAATTAA	AATTGTAAAT	CAATTTTAGA	168060
AAAAAATTAT	TTACACTATA	ATATAATTTT	ATATGAATCT	AGGGAAAAAC	AATCCAAATA	168120
TAATTAAAAAT	TGGTAAAATA	TTTAAAAAAA	ATAACTACGA	ATTTTATTTA	GTTGGAGGCG	168180
CTTTAAGAGA	CTTACTGCTT	AATAAACAGC	CTTACGATTT	TGATTTTGCA	ACAAATGCAA	168240
CTCCTGAAGA	AATAATAACA	TTATTTCCAA	ATAACATCAA	AACAGGAATA	AAACATGGCA	168300
CAATTGGTAT	TATTTTTAAT	AAAAAATCT	TTGAAATCAC	CACATACAGA	ATAGAAAAAG	168360
AATATGAAAA	CAACAGAGCC	CCCAAACAAG	TAGAATATAC	TAAAAATTTA	CTTAAAGATC	168420
TTGAAAGAAG	AGATTTTACA	ATTAATGCAA	TTGCAATGGA	TATTTTCAAC	TTCAACATAA	168480
TAGATTGCTA	TAATGGGAAA	AAAGACCTTA	ATAAGAAAAT	AATAAGATGC	ATAGGAAATC	168540
CAAACAAAAG	ACTTGAAGAA	GACGCCCTTA	GAATACTTAG	AGCAGCAAGA	TTTTTCATCCA	168600
CACTTAATTT	TAACATTGAA	AAAAATACTT	TAATTTCAAT	GAAATATAAA	AAAGAAAATA	168660
TTTTAATGAT	TTCAAAAGAA	AGAATAAAAA	ATGAATTTCA	CAAATTGTTA	GAAGGCATAA	168720
ATATACAAAA	AGGAATTTAT	TATCTTAAAA	AAGTTGATTT	TTTTAAAAAT	TTTTTTAATC	168780
TAGAAATAAA	AACAAAAGTA	ATCAAAAAAA	TTGCTCTACT	TGATAAAAAAC	AAATTTTATC	168840
TAAAGGCAAT	CACAATATTG	ACAATTAAAA	AACCTATAAA	AGAACTAAAA	GAAAAATTAA	168900
CTTTACTTAA	ATTCTCAAAT	AAAGAAATTA	AGCTGATTTT	ATTTTATAGA	GGCATAATCG	168960
ATAATAACAA	TATTTTTAAT	GTCAAAAAAT	TAAGTGATAT	TAGATATTTG	CTTAGCAAAA	169020
GCACAAGAGA	ACATTATAAA	GAAATAATTG	ATATATACAA	AGCACTCAAA	GGAAAAAATA	169080

AAAGATATTT	ATTTATAATA	AAAAACATAA	AAAGAAAAAA	ATTGCTAAAA	ATTCCTCTCT	169140
CTTTAAAAGA	TTTAAAAATA	AACGGAAAAG	ATATTCAAAA	TCTAGAACAA	ATAGAAAACA	169200
AAAATATAGG	TAAAATTTTA	AATATGCTAC	TAAGATGTGT	AATTGAAAAT	CCCAAGCTTA	169260
ATACTAAAAA	TTATCTTATA	AAAAAATCA	AAACCTTAAA	GGTTAATGTT	TTCCATAGCT	169320
TTTAAAGCTA	CTTCAGCCGC	TCTCATTTTCG	GCTTCTTTTT	TAGATTTGCC	CTTTCCATTT	169380
GATATAAAAT	TTTCTCCAAC	ATAAAGTTCC	ACACAAAAAA	CTTTATCATG	GTCTGGACCT	169440
ATTTCTTGT	CTAGCTTATA	ACTTGGCGAG	ATTTTATATT	TCTTTTGAAC	ATATTCTTGC	169500
AACAACTCT	TATAATCTTT	AAAATCCCCC	CTATTAAACA	TCAATCTTAT	ATACATATCA	169560
AAAAGTCCAA	CCACAAATTC	TGTTGCTCTT	GAAAACCCAC	TATCAAGATA	AATAGCGCCT	169620
ACAAAAGCTT	CAATAGCATC	TGCAAGAATG	CCTTTTTTAT	TTGACCATC	ATTACTCTCC	169680
TCCCCTCTAC	CTAGCAAAAT	ATAAGAACCA	AGATTAATCT	CTCTAGCAAT	ATTAGATAGG	169740
GAATCTTCAC	TAACAATATA	AGATCTGGCC	TTACTGAGCT	CTCCTTCACT	TTTATTGGA	169800
TAAGTTTTAT	AAAGATGATC	TGTAATAATC	AAATTAAGCA	CAGAATCTCC	CAAAAATTCT	169860
AATCTCTCAT	TATTACTAGA	TTTTTGATCC	AACTCATTAG	AATACGACGA	ATGACACAAT	169920
GCTGTATTCA	ATAAATCAAA	ATTACTAAAG	TCAATGCTCA	AATTTTCCAA	AAATTTACTC	169980
AATTGAGATT	TTCTTTTATT	ACACAAACAA	AAATCAGAAG	ATTTTTTTTT	CATCAACCCT	170040
TTCTCTTTTT	AATAAAATTA	ACAACATCGC	CTACCGTCTC	AAATTCATTG	GCTTCATTCT	170100
CTGGAATCTT	ATCATCAAAG	GCCTCTTCAA	GCAAATACAA	AAGCTCATAA	ATATCTAGAC	170160
TATCTGCATT	AAGATCTTCA	ACAAATCTAG	AGTCTGTGGT	AATTTTCATCT	TCTTTTTTAT	170220
CAAGTTGCTC	AGATATAATA	GACCTAACCT	TGCTAAAAAT	TTCATCATTA	TCCATGAATA	170280
CACCTTCCTT	ACTACAAGCT	ACAAACCTAT	TTCTAGATAT	TGGTTATTCC	TATAATATCC	170340
ACATTTTAAA	CAAATCCTAT	GTCTCACGCC	AAGATTACCA	CAATTAGAAC	ATTCTTGAAA	170400
TTGTGGAATT	TTTTTTCTCA	TATTTATACT	CCGCCTTGTT	CTACTTCTAG	ATTTTGAAGG	170460
CTTAAATTTT	GGAACAGCCA	TTACTTTCTC	CTAACTACTT	TTAAACTACA	TTAAATTATC	170520
TTAATAATAT	ATATATTAAC	CAAGTCATTT	GTCAATAAAC	TTAGATTTTA	ATCTATTAAA	170580
CACCAATTCT	GGAACAAAAT	TAGAAAGATC	AACATCCTTT	TTCAACATCA	ATTCCTTTAC	170640
AAAATCCGAC	CTTACATATA	AATGTTCTGC	ACTACTTGGT	AAAAATATAG	TATCAATTTT	170700
AAAATTTAAC	TTATTATTAA	CAAGATATCT	TTCAAACTCT	ATATCAAAAT	CATTAAAAGC	170760
CCTAATTCCT	CTAACAATAA	ATTTAATAGA	ATTAATTAAT	GCATAATCAA	CAATAAACCC	170820
GCTATACCTA	TCTACAAGCA	CATTTGAAAA	ATTTAAAGAC	GAAATAACAT	CTTTTGTAAG	170880

GCTAAACCTC	TCAATATCAC	TTAGGAAATA	TTTTTTTGAT	TTATTTTGTAG	CTACTAAAAC	170940
AATAACTTTG	TCAAAAATAG	CCAACGATCT	TTTAATTAAA	TCAATATGAC	CCCAAGTAAT	171000
TGGATCAAAA	GATCCTGGAA	AAACTGCCAC	CCTCATATCA	AGAAAGCCTA	AACCCCCTAT	171060
TAATATCTTT	GAGAATTTTT	TTACGCCATA	GTGGTTGTTC	TTCTAAATTT	TCAAATTCAA	171120
TTTCTTTAAT	AATTTTGAAA	TATAAATATT	CTCCTCTAGA	AGATGTATCC	ACAATAAAAA	171180
ATTTACCAGC	ACCTTTATAA	ACTACTTTTT	CATTACTTGA	AAGATTATTT	TCAGATGTTA	171240
ATTTATTGAT	AACACAATCA	AAATGAACAG	GTTTATCTTC	ATTTTCAACC	TTTAAAGACA	171300
TGCTATAAAC	AATATCATTT	ATTTTTTTTT	CACAAATAGC	ACAAGCAACA	TCAAGATTTA	171360
TTCTTGTTTT	AAATTTTATA	CTTTTTCTTC	TAAATTTACC	TTTAAAAACA	TTTTGTTTAT	171420
CTGTATTCAA	GACAGTCGAT	TTATTATTAA	AACTGTTTAG	CGGGGCAACA	TTTTGACTTT	171480
TGCTCTCAGA	ATTTAATCTT	TTATTCCTAA	AACCATAAAA	ACGCTTATTG	TGTTTCTGAA	171540
AATTACTACT	CTTGAACTTA	CTTTGAGTAT	ATTTCAAAAT	TAACCTCTCT	TAACATATAA	171600
AAATAAATAT	TTACTCTTCA	AAAAAATAAA	AACAAATATT	TAAGAACAAA	ATCACTCACA	171660
TAGCCTGAAA	GTTTTTCAAA	AAATCAAAAC	TTTAAATCTT	CAGCTTTTTA	AAGACAAAAG	171720
CACAATTAAA	CCCTCATTGA	AATTTAATAA	GAATGGTTAA	TTAATGCTCT	CAATTATAAT	171780
AAATACACAA	CCAAGCTTGC	CAAAATTATT	TCCCAAAACA	GCAAAACTC	TACTAAATCT	171840
TGAATAAAAA	TTTACAAAAC	AAAAGATCAA	TCAAAAACAA	TTTTTAACCT	CAATCTAATT	171900
TGATTCATCT	TATGACCTAA	TTTGCTAAAA	TATTTAAAAT	AATTAAAAAC	TAATGCTTAA	171960
CTTTTTTAAT	AGTAAGGCGC	TCCTTTATCT	TCATAGCAGC	CTTACCGATT	CTATTCCTCA	172020
TATAATAAAG	CTTTGCCCTT	CTAACTTTTC	CCCTTCTTAA	AACTTCAACC	TTTTCTATAA	172080
TAGGAGAATA	TACTGGGAAA	ATTTTTTTCAA	CACCTATTCC	TGAAGAAATT	TTTCTAATCA	172140
AAAATGTTTT	GCCAATTCCC	TTGTTTTGGA	AAGAAATAAC	AATCCCTTCA	AAACTCTGCA	172200
ACCTTTTCATT	ACTACCCTCA	ATAATTTTGT	AAACAACCCT	CACAGTATCT	CCCACATTAA	172260
AAACAAAAGC	CTCATTTTTTC	TTATTC TGAG	CTTCAATTTT	TCTTATCAAA	TCCATTATCT	172320
TCTCCTATTA	TCTCTAAATA	TTTAAGGTAT	AAATCATATC	TATTTTTCTT	AGTTTTTTCT	172380
CTAGCTTTAA	CAAGCCTCCA	ATTCTTTATA	TTTGCATGAT	GTCCCGAAAG	AAGAACTTCT	172440
GGGACCTTTA	TCCCCTTAAA	ATCATAGGGC	CTGGTATAAT	GAGGATATTC	AAGCAATCCA	172500
TTTTTTTACAC	CAAATGATTC	TTCTAATAAA	GAATTGGGAT	TTATTACTCC	ATCTAGCAAC	172560
CTATATACAC	TATCTATTAA	AACAAGAGCT	GCAATCTCTC	CTGAAGATAA	AACATAATCT	172620

CCAATAGAAA	TCTCAAAATC	ATCATACAAA	TCTATAATAC	GTTGATCAAT	CCCTTCATAT	172680
CTTCCACAAA	TTATAACAAT	TTCTTCTCTT	TTTGACAAGG	AATACGCCAA	CTCTTGGCTA	172740
TACTTTATCC	CAGAAGGACT	TAAAAATATT	GTCGTTTTCT	TGGCAGACTC	TACATGCTCA	172800
AGAGCAAAAG	AAATCGGTTT	GGCCTTCAAT	ACCATCCCAG	CACCGCCTCC	ATAAGGCAAA	172860
TCATCACATC	TTTTATGCTT	ATCTTTTGAA	AAATCTCTAA	CATCAACAAG	CTCAAAACTT	172920
ACTATTCCTT	TATTAATAGC	TTTTTTTCAAT	ATTGAATTTT	CAAAAAATGG	CTTAATTATT	172980
GCTGGAAAAA	GGGATAAAAC	CGTAAATTTT	ATTTTAAAAG	ATCTAGAACC	TTAAGCTCAA	173040
TTGTTTTTTC	TTGAGTATTT	ATATCTCCAA	TATATATACT	TAAAAAGGGA	ATAAGAAAAA	173100
ATTTAATACC	CACTCTGACC	TCAAGAAATA	CACTATTTAA	ATATTCAAAG	AAAGCTACAA	173160
CTTCTCCTAG	TTTTTTTATTA	TTATTAACAA	TGGCATAGCC	AATAAGCTTT	CCTAAATAAT	173220
ATTGCGCTTC	TTTTTAACTC	GATGCAAGCG	AATCATCAAC	CCACAATTCA	AAACCAATCA	173280
GCGGCCTAAC	TGCCTCTGGA	GTATCAATCT	CTTCAAACCT	CAAAAAATAAG	GAATTACCCT	173340
TTATATTAAC	ATCTACAAC	TTAACTTCAA	CACTGGAACT	ATTGCTTTTT	TTTAAAAGAA	173400
CTTTATTGTT	TTTTAGATTA	ATAAAATCAC	AAAAATTATT	GGATATGCTT	TTAACCCTAG	173460
CATACCCATT	AACCTCATAA	GACGATAATA	TAACGCCTTT	AATAAACATA	GATTAATCTA	173520
AAATTTCCAA	TTGCACTCGC	CTATTGGTTT	TGGCAGCACA	AGCTCCAAGC	AAAGTTCTAA	173580
TAGCCCGCGC	AATACGACCC	CGTCTTCCGA	TTATCTTGCC	CACATCACTT	TGAGAAACCC	173640
TTAATTCCAA	AATAGTTGAT	TTTTCCCCTT	CAATTACATT	TAACTTTACT	TCATCTTCTT	173700
TATCTACAAG	AGACTTTACT	ATAAACTCTA	TAAGTTCAAT	CTCATTCCCA	TACTCTTTCA	173760
TTTAAACCTC	CTGACTTTTC	GCATTCAAGT	TGTTTTTATT	TAAAAGCATT	TTCACTGTAT	173820
CACTTAAAT	TGCTCCCTTG	CTTATCCAAT	CTTTCATTCT	ATCTTCCTTA	ATTTTTATTT	173880
GGTTTTGCTT	TTCAACAGGA	TGATAATAAC	CAAGTTCTTC	AATTGCTCTA	CCATCTCTAG	173940
GAGACGTAGA	ATTCATAACT	ACAACCCAT	AATAAGGTCT	TTTTTTAGCT	CCCATTCTTT	174000
TCAATCTTAT	CTTAACGCTC	AAATTTATTC	CTCCTTATTT	TCCCAAAAGG	GATGCAATCT	174060
TATTTTGAAA	ATCTTTATTT	TTCATTTTTT	TCATAATCAA	AGTTGCTTGA	CTAAACTTTT	174120
TTATGAGCTT	ATTAACATCA	AAAACAGTTG	TTCCACTTCC	CATGGCTATT	CTTTTTTTTT	174180
TTGAGGGATT	ATTCAAAATC	ACTGGATTTA	TTCTTTCTTT	TTTAGTCATA	GAAAGAATAA	174240
TAGCTTCTTC	TTTATTAAAA	CTTCTTCAT	TTAAATTATT	GCTATTCAGC	ATTGATTTTG	174300
AAACACCTGG	TAAAAAATTT	ACAAAATTAG	AAAACCCTCC	TACTTGCTTA	ATGCGCCTAA	174360
ATTGACTCAG	ATAATCTTCA	AAATTAAAAC	TGGCTTTATT	AATTTTTTCT	TCAAGCTTAA	174420

TAGCCTCTTC	TTTGTCAACA	ACACTTTGAA	CCTTCTCTAC	AAGACTAACA	ACATCCCCCA	174480
TGCCAAGAAT	TCTAGAAGCA	ATTCTTTCTG	GGTAAAAGGA	ATCAAGATCT	TCGATTTTCT	174540
CTCCAACACC	AATAAATTTA	ATGGGAACTG	CACAAATACT	TTTAAAAGAT	AATACAGCTC	174600
CCCCCCTAGT	ATCTGAATCA	AACTTAGAAA	ATATTGCACC	GGTAAGTCCA	ACATTCTCAT	174660
TAAATTCCTT	AGCAATATTT	ACAGCAACTT	GCCCCATCAT	AGAGTCTACT	ACTAAAATGG	174720
TTTCTGCGGG	TCGCAAAATC	CCCTTTATTT	TTTTTATCTC	TTCAACCAAC	AAAGATTCAA	174780
TTTCAAGTCG	TCCTCTAGTA	TCAACTATTA	CAGAATCAAA	AAAATTAGAT	TCAGCAAAC	174840
TCATAGACGC	TTTAACAATT	TTAATAGGAT	CTTTTCCCC	TTCAATTGAA	AATACTGGAA	174900
CACCTACTTG	ACCACCCAAT	ATTTTAACT	GTTCTACGGC	CGCCGCTCTA	AAAGTATCAG	174960
CAGCTACAAG	AAGTACTTTT	CTATTTTCCT	TTTTAAGCTT	TAAAGAAAGC	TTGGCGCATG	175020
TCGTGGTCTT	GCCAGAACCT	TGAAGTCCCA	ACATAAGAAT	ATAAGATTGC	TTATTGGCAG	175080
GATGTAAACT	AAGCTCATAA	TTTTTGCTC	CCAAAAATT	AACAAGATTA	TCATTGACAA	175140
TTTTAATAAA	CTGAGATTTA	GGATCAATGC	CCCTTAAAAC	TTTTACTCCC	TTGGATTCTT	175200
CAATTATAGA	ATTTAAAAAA	CGCCTTATAA	CTCTTAAGTT	AACATCAGCA	TCAACTAAAG	175260
AATTTTTAAT	AATCTCAATA	GCCTCTGCAA	TGTTTTTATC	ATTTATCGTA	GATTTTCCAG	175320
AAAGATAGTT	TATAAAATTT	CTAAAATTTG	ACCCCAAAC	TTCAAGCATT	AAAAACAACC	175380
CTTATTCTCA	AAAATAATCT	TGTCAAATAT	AAGATACAAA	AAAAGGCAAT	ATTAATCAAT	175440
ATCTCCTCCA	GTCAAATAAA	ATTTATTAAA	AATAACACTG	GAAGCGGGCC	CAACAACAGA	175500
AACTTCAGTA	TCTAATGAAC	TAAGCTTTAA	TACTATATTC	TCTTTATTGA	GCTTTTTAAT	175560
CTCTTCTTTT	AACAAATCAA	AAAAAGCTTT	TAATTTAAAA	CTTGACCAT	AGAGTACTAA	175620
ATAATTAAAA	TCAAGCATTC	TTTGAATATT	AATAATAATT	ATTGCCAAAT	ATTTAACAGT	175680
ATCTTGATA	ATTTTATTTA	TAAAATCATA	TTTTTCATAA	AGAGAAAAAA	TGTCATATAT	175740
TGTAACCTTT	TTCAATCTGC	CCTCATACTT	TTCATAAAGC	TCAGGAATCT	CACCATTTCAT	175800
AAATTCCTTG	GAAATCAATC	TCTGCAAAGC	AAAATTAGAT	ATTAGCATAT	TGACACAACC	175860
CTTATTACCA	CAAGTTGGAC	AATTTTTTTC	TCCCTCATAA	TCAATTATCA	TGTGACTAAC	175920
CATACCTGAT	TTATTATTAA	AACCAGGGTA	AACATTGCCG	CCTGACCAAA	TCGAAAGTTC	175980
AGCAGTATCT	GTGTAGTCAA	AAAACATAAT	ATTATCTATA	TTTTTACCCA	TAAATTCAGC	176040
AAGAGATAAA	TTTTTAACAT	AACTTTCAAG	ATAAACTGTT	AGTGAAAAGT	ATTCCTCAAG	176100
TATTCCTTTA	ACAGGAACAT	CTTTTTCAAT	CCACGATCCA	TAGCTATCAT	TAACAATGCC	176160

CAATTCTTTA	TCCTTTATTA	ACCTGTAAT	ACTAAAGCCT	AAGCCAATAA	ACTTATCTCT	176220
TGAGAAATTA	TGTTTCCAAA	TAATTTCTAT	CATATGATCT	TTTATTTTTT	CTAAAAATATC	176280
ATAAGCGCTA	ACTGGGGGCT	CAAAAGAATG	AGTCTCGCTT	ATTAAAACCT	CGCATTTAAG	176340
ATTGGCAATA	CCTATTTGAA	AATAATTGCT	AGAAATAATA	ACTCCCATTG	AATACGCATA	176400
ATCTTTATTA	ATATCGAGAA	GTATTTCTTT	TCGTCCATGT	TTTTTAACAT	CAGACACCCT	176460
AGAGCCAACT	TCAATCAAAA	GATTTTCTTT	TATCATTTGA	TTAGTCAAAA	TAGTAACTGC	176520
AGCATTTGTC	AAGCTTAACT	TACGAGCCAG	GTCTGTTCTT	GAATATTGCA	TATTTTTTCAA	176580
ACTAAGAAGA	ATTTTTCTTC	TATTTCCGCC	TCTAATTGAA	ACCATATTTT	CACCCTGCAT	176640
AATACACCTC	CTTTATTTTT	AAAATAAAAA	AAATATTATA	AAATATCATA	TCAAAAAAAC	176700
CAATACAATA	TTTTATCTAT	TCTTAAAAGA	CAAACATGCC	TTTATAAGGC	TAAAAAACAT	176760
TTTACATCAT	AATATCACAT	TCATAAAATA	TCAAAAACCT	AAAGCTTAAC	AAAAAAGGGA	176820
ATAAAATCAT	TTTTACATAA	AAACTCATCA	ATAAAATTAA	TTGGATTTAA	AAATAATAAA	176880
TACAAGAAAA	GCCATTTTGC	CTTAAAAACC	ACTAACTTTA	ACTTAATTTT	TCCTTAAAT	176940
AAGAAAATTC	CATAGTAAAA	CTGCCCCCTC	CTTTAGTAGA	ACTTCTCAAA	ATAGAAGCAT	177000
ACCCAAAAAG	TTTCTCAAAC	GCCGCCTCTG	ATTTTATCAA	GTCATACTCT	CCAATATTGC	177060
TAACTGAATG	AATAACACCC	CCCATAACAT	TTAATGTAGA	AATAATTTTCG	CCTGTATGCT	177120
CAATGGGTGT	TCTAATTTCT	AATAACATTA	TTGGCTCAAG	TCTAATAGGA	TCTGATTTTT	177180
GAAAAATACT	ATGAAAAGCA	AATCCTGAAA	TTGACTCAAA	AGCACTCTCG	CTAATCTTAT	177240
TGGCCCCACA	AACAATAGAA	AAAATACTAA	CATTAATATC	AATAATGGGA	TATCCAAAAA	177300
TTCCACTTAC	AAATGCAGAT	GTAATTCCCC	TCAATATTGC	AGACTTAATT	ACAGGTTCAA	177360
TGCCACATTC	AAAATCAATT	TTATTTCCCG	CACCCCGCGA	CAAAGGTTTA	ATGATCATTC	177420
CAATTTTAAA	ATCAATATTT	TTGCCAGCAA	AAATATTGTT	AAACTCAAAA	ACTTCTTTTA	177480
CAATTTTGCC	AGCACTTTCT	CTGTAACCTA	CTTGAGGCTT	TCCTGTATAA	ACATTAAGAT	177540
TAAATTCATC	TTTAATTCTT	GTAAAATAA	TCTCAAGATG	TAATTCACCC	ATTCCAGATA	177600
TAATTAATTG	CCCTGTTTCT	TTACTCTCAG	AATAACTAAA	GGTAGGATCT	TCTTTAGATA	177660
TTATTTCAAA	AATTTCCCTA	AGCCTAACCT	CATCTGATGA	TCTTTCAGGC	TCAACAGACA	177720
TTAAAACAAC	CGGCTCTGGA	AACATAACAG	CCTCAAGTAA	AACATTATTA	TTCTCTTCAA	177780
CAAGAGTATC	TCCTGTAACA	GAAAACCTTA	ATCCCAAAAC	AGCACCAATA	TCGCCTGTTT	177840
TTACAAAATC	TATTTGTTCA	TTTTTATTTG	AAAAAACTCT	AAAAATTTTT	GTAAACTTTT	177900
CACGCTTACC	ATTTGAAGCA	TTGATAATTT	TTTTATTAGG	ATTAATCTCG	CCAGAATAAA	177960

CTCTAACAAA	ATAAAGATGA	GAAGCAATCA	CGCTTGAATA	TTGAACTTTA	AAAACAAGGG	178020
CTGACAATTT	TTTATTTTCA	TTAGGATCAA	CTAAAATTTT	TTTATTTGTG	TCTAAAGAAA	178080
AAGCGCTAAA	ACTTTTTTCA	AAAGGACTTG	GCAAGTAATC	TACAATCGAA	TCAATCAAAG	178140
GTTCTATTCC	AATATTTTTT	AAACTAGTTC	CCATTAAAAC	AGGAATAATA	AATCTAGAAA	178200
TAGTGCCTCT	TCTAATCTCT	CTTTTAATAA	TATCTAAACC	AATCTCTTTG	TCTTCAAGAA	178260
ATAATTGAGT	AATTTCTTCA	CTAAATTGGC	TAAGAATGTC	TATTAATTTT	TTTTTAAAAA	178320
GAATCACTTT	TTCAATAAAT	TCTTCTCTAA	TTTGACTATA	AGTTAATTTT	GGAATTCCAT	178380
TTTCCATTGA	AAAATGAAGC	TCTTTATTTA	AAATAATATC	AACTACTCCT	TCAAAATTGC	178440
TTTCATTTCC	AATTGGAATT	TGCAAAACCA	AAGGAATAGT	TTTAAATTTA	TTTTCAATAT	178500
CTCCCACAAC	TTTAAAAAAA	TCAGCACCTA	TTCTATCCAT	CTTATTAACA	TAAGCAAGTC	178560
GTGGGATTTC	GTATTTTTCT	GCCTGTTTCC	ATACAGTTTC	TGTTTGGGCC	TGAATTCCAT	178620
CAACAGCGCT	AAAAATAACA	ATACCCCCAT	CAAGAACTCG	AAGAGATCTT	TCAACTTCTG	178680
CTGTAAAATC	CACATGCCCA	GGAGTATCAA	TAATGTTTAT	TTGGCAATCC	TTCCAATGAC	178740
AAGTAATGGC	AGCTGAACTA	ATAGTAATTC	CTCTTCTTTG	CTCTTGAGGC	ATCCAGTCAG	178800
TAATAGTGTT	TCCTGAATCT	ACATCCCCCA	TTTTATGACT	TTTGCCAGTA	TAATAAATAA	178860
TTCTTCTGT	GGTAGTAGTT	TTTCCAGCGT	CAATATGAGC	CATAATTCCA	ATATTTCTAA	178920
TACTCATAAA	TCCCCAACAA	CTACCACAGC	CTCAATGCAG	ATAAATTCAT	TAGCACAGGT	178980
TAAACCTTTT	AATAAAATTG	CTATTTTGCA	TCACACTCGC	AACATTTGCT	TTCATTCTCA	179040
TCGTGACAAA	AACAATCGCA	ATTTTCAAGC	AATGCCTTAT	GCATCCATAA	ATGTTTCTCA	179100
AGATCACTCA	TGATATCATC	CATAATATTA	GCAGTACCAT	AATCACCAGC	AGTATCAATT	179160
AATTTTCTCA	TTCCAAAAAT	ATTCTTCAAA	ATCTCAGTAA	GACTGCAAAC	AATGCTTTCC	179220
ATTGAGGGCA	AAAAATTAGA	AGTTGATTCA	ATATCAAGCT	CCTTAATAAA	GGATTTTTTC	179280
ATAAACTCAG	AATATCTAAA	TTCAGAATCA	TATCCAAGCA	TTCTTGAGCG	TTCTGCAACA	179340
ATATCAATAA	TTTTTTCAAT	ATATTCATAA	AGTTTTTTGAG	TTTTTTTGTG	AATAACAAAG	179400
AAATTGGTAT	CTTTTATATT	CCAATGAATA	CCTCTTAAAT	TAGAATAAAA	AATATGCAAA	179460
CTTGCTAACA	ATTCTTGTA	TTTTAATTGT	ATTGCGTCTA	AATCATCCTT	TTTTATATAG	179520
CTTAAATACT	TTTCCATAAC	TATCTCCTTT	ATATAATTAT	TATAATACAT	AATGAGATAT	179580
AATTATGGTT	TTAATACCAT	AATAAATAAA	AAGGATATTT	AATGAAAAAA	TTGATTATAA	179640
TTTTTACACT	GTTTTTATCT	CAAGCATGCA	ATTTAAGTAC	AATGCATAAA	ATAGATACAA	179700

AAGAAGATAT	GAAAAATTCTA	TATTCAGAAA	TTGCTGAATT	GAGAAAAAAA	TTAAATCTAA	179760
ACCATCTAGA	AATAGATGAT	ACCCTTGAAA	AAGTTGCAAA	AGAATATGCC	ATTAAACTGG	179820
GAGAAAATAG	AACAATAACT	CACACCCTTT	TTGGCACAAC	CCCAATGCAA	AGAATACATA	179880
AATACGATCA	ATCCTTTAAT	TTAACAAGAG	AAATACTGGC	ATCAGGAATT	GAACTTAACA	179940
GAGTAGTTAA	TGCATGGCTT	AATAGTCCAA	GCCACAAAGA	AGCTCTTATT	AATACAGATA	180000
CCGATAAAAT	AGGTGGCTAT	AGATTAAAAA	CGACTGACAA	TATAGATATA	TTTGTTAGTTC	180060
TTTTTTGGAAA	AAGAAAATAT	AAGAATTGAC	ACCATTAAAG	CTTATACTGT	ATACTACTTA	180120
TTAGTAATAA	AAGGGCTCAT	AGCTCAGTTG	GTCAGAGCGC	CTGCCTTACA	AGCAGGATGT	180180
CGGGAGTTTC	AATCTCTCTG	GGCCCAAAAA	TAATCTAAGT	CTCAATTACC	TTTAGCTTTA	180240
AAAGCAAATC	TATTTTAGAA	TCTTTAAGCA	TGTTATTTAA	TTCTTTTTGC	ACTATATTAG	180300
CTTTTATTAG	GCTTTCAAAA	AGAAAAATAA	ATGCTCCTCC	TTTACCAGCG	CCACTTAACT	180360
TACCAGAAAAG	AGCGCCCAAT	TTGATTCCCT	CACTTATCAG	CCAATCAAGA	GTATCATTAG	180420
ACAACCCCAA	ACGCTTTAAA	CAACATTGTG	CAATATTCAT	TTCATTAGCT	AAAGAATACA	180480
CATCCTTATT	CTGAAAAGAA	GCATAAGAAT	TGCTTACGGC	AAGGCCAAGC	TTTTCAATAA	180540
AAACAAATAA	ATAAGCATTT	GATAATAGAT	CTTTTTTCAA	ATTAACAACT	ATTTCTTTAG	180600
TTGTAAATC	TCTTTTTATT	GCTCCTATTA	GAAAATAAAA	ACCAGAATCT	TTTATTTTCT	180660
TTGAATGTAA	AACATTTTCT	TTTTTCTCTA	AATAAAAAGT	TCCATTAAGA	TCGATTAGTC	180720
TAATATCCAT	TCCAGAAGAT	TTGCCATGAA	AAATGTTTTT	AATTTGATTT	GCCAACAAAA	180780
TTTTATTACA	ATCCTTATAT	TCAAAATGAC	TTGTAATATA	TTCTGCAAAG	CATAAACTAA	180840
GACTAGCAGA	AGAACCAAGA	CCAACTCCAA	TAGGAATTTT	AGAAATTATA	TCAAACCTCAA	180900
TAGGATTAAC	TTTGCTATAA	TTTGAAACAA	TAAAACTTAT	AAGGCTATTT	AATCTTGATC	180960
TGGGTTTTCC	TAAATATTTT	CAATTTTTAG	ATACACTATA	AATCAGATCC	ATATAAATTG	181020
GAAGTGTAGC	GCCAATAACT	GGGAACCCAT	AAACAGCGCT	ATGTTGCGCT	AAGAACAATA	181080
TTTTAGCAGG	CTTCTTATT	CTTAGCATTT	ATCGCTTTCA	AACTTAATGC	CTTCATTCTC	181140
TAAAACAATT	GAAATAATTT	TTGACAAATT	AAAAGCTTCA	ATATTGGGCC	TGTAAACTAA	181200
AAAAGTTTCG	TTTCCGGCTC	CCAAAGCTTT	AATCAAATCA	CATTGACCTA	AAAGGTGATC	181260
AAAAGTTGAA	GGCAAAGCTG	CCGAAACTCC	TATTGCCTCT	CCAATTGCCA	ATCCCAATTC	181320
TTTAGCCCTT	CTTAAACTAG	AAATTAATGC	GGATTTGGAA	TTGCTAGCAT	TTAAAACAAG	181380
CTTTTTTCATC	TCCAAATTGC	ATTTCAATAT	AAAATCTAAA	ATAGAATTTT	TATGTTTATT	181440
GTATTCACAA	ATAGACGTAG	TAGTTTTAAT	TGCTTGCAAA	CCCTGCATCA	AATAAAAATC	181500

ATTAAACTCT	ACAGCACCCA	GCTGCCTGCA	TTTAGGATTA	AAGCCGCCTT	CAAACTCAAT	181560
AACACCCCCA	AAAATACTAG	TAGCAATATC	ATATCCACTG	CCTATTCCTC	CTTGAGAATA	181620
CCTGTAAGCT	TCCAAACAAT	ATTTAAAAAT	TTCACCTTTC	TCAACAACAT	TGGTAGCATT	181680
GTGAATTAAA	AAAAGCCCAC	ACACTATGCC	AATAGCAACG	ACAGCACTTG	AACCAAATCC	181740
CTTTTTAGTT	CCATCATTA	AGAAAAAATT	ACTTGTATCA	ATATATACAT	CATACGCAAA	181800
ATTCTCTAGA	TTAAAAAAC	AATTTTGA	CAAGTAAGCA	AACATTTTAA	AAACAAAATC	181860
GCTTCTATTT	TCTATTAAAG	AAAAATCGTC	TATTTTTTTC	TTTTTACTAA	AAAAGCGCCA	181920
AGAATCGCTC	TTTTTAAAG	AAAAAATGC	TCTCTTGTG	ATGGCAATTG	CCAACCCCAA	181980
TCCCTTTTCC	TCTAAATAG	TATACTCCCC	CATTAAGT	AAATTTCCGG	GTACAGAAAA	182040
ACTAATCAAA	TCCATTCTAA	GTCACACCCA	ACCTTTGAAA	CAATAAAATC	AATGCCAGTA	182100
AAATTTTGCT	TAAGTCCTTT	TAAAATAGTA	TTTAAATTTT	CCTCCAAACA	AAGAACTTT	182160
ACTTGGGGGC	CTGCATCCAT	CGTCTCAAAT	ACAAAATCC	CCTCATTTCT	CAAATCAGCA	182220
GCATACCTAA	TTAAATCTAT	TGTACTATTT	TTAAAATAAA	AAATAGAAGA	TGCAAACATT	182280
AAGGCAAACA	TATTCTGATA	ACTTTTAC	ATAGTTGCTC	CAAAATGTAT	AAAATCCTTT	182340
TTTTAAAAAA	AAATATAAAG	CGTCTTTAAA	AATCTTTT	CTAGAGGCAA	TCCAAGCATC	182400
ATAATAAAAT	TTATGCCGTT	TGCAAAATAT	CATTGCGGCT	CTTGAAGACA	ATTCTTTTTC	182460
ATTACTATCA	ATTATGGCAA	ATATTATTCG	CAAATCATTA	AAATAAGATT	GATCTCTTAA	182520
TTGAAAAGAT	TCTTTTGAAC	CCTCTTTTAA	AATAGTAAAC	CCCCCGTAAA	TAGCCCTTGC	182580
CGCAGAAGCC	GATCCTACTC	TTGCAAGATT	AGATGCGCTA	TTACAAGAAT	ATTTATTAAA	182640
ATATTTCAAA	ATACAAGCAG	CAATAGAAGC	AAAACCTGAA	CTTGAACCTG	CAAGGCCTGC	182700
TGCTGTGCGG	AAATTGTTTT	TACTTTTAAT	TTTAAATCTA	ACATTTGGTT	CATTAAGAAT	182760
TTTTCTTGCA	TAATCAAAAA	ACACCTTTTC	TCTATTTTTT	AATATAACTG	GCTTTGAATT	182820
TAAAATTATT	TCATCTCGAT	TTGAAAGTTC	AAGCTCAC	ATTGAATAAA	ACTTGTCAAC	182880
ACTAACAGCA	AGACTGGAAG	TAGCTGGAAT	GTTTAAAAAA	ACATCCTTTT	TCCCCCAATA	182940
TTTAATTAAA	GCTAAGCTTG	CATGAACTTT	ACACTTTATT	TTCACTCTCT	AACCTTATTT	183000
TCTTCAAAAT	TTTAAAAGCA	AAATCAAAAG	AATAAATATT	CATTCTTTCC	ATTTCCAATA	183060
ATAATTTGTC	TTTTTCAAAA	TCAGAAATAT	TATATTTTGT	CTTTAAAAGA	TGGAGTATTT	183120
TATTAACATG	CAATCTCATA	TGACCCTTTT	GAATCCCAT	AAATGCAAGA	GCCCTTAATG	183180
CTGCAAAATT	ACTAGCAAGT	CCAACACAAG	AGAGAATACC	AATAAATTCG	CTCTTACTAT	183240

TTACATTCAT	AATTTTAAAA	CTTAAAATTG	AAGCTTCATT	AAAAGATATA	ATCCCCACCTT	183300
TAGTTCCAAC	TTGCAAAGGA	ATTTCAATTT	CTCCAACATA	AGCATTGTCA	GATGATATAAA	183360
ATTTACTAAG	GGGAAAATAT	TTGCCGCTTT	TTGAAGCAAA	TTTATGAACA	GAGGCCTCAA	183420
GAGCTCTTGT	GTCATTAAAA	GTTGCAAGAC	ACACCCCTGT	AATTCCATTC	ATAATACCTT	183480
TATTATTAGT	AACAGCTCGC	TCCTCTTCGT	AAAAACCTAT	ACTAGAAATA	AGTTCAATTT	183540
TTTTAGCCAA	ATTCCAAGAA	TCCTCTTTAC	CCGGTAGCAA	ATGCTTAAAA	TCTAAAACAA	183600
AACGGGCTTT	GGCTGTAAAT	TCACTAATAT	CATTGCTTAA	AACTTTTAAA	ACACACTCAT	183660
ATCCGAATTC	TAAAAAATA	AATTCTGCTA	CACGCTCTGC	AATTGAGTTT	AGCAAATTAG	183720
CACCCATAGC	ATCACAAGTA	TCCACATAAA	TATTTAATTT	TTGAATACCT	AATTCTTTAA	183780
TGTGCCTAGT	TGACAACCTT	CTAAATCCAC	CCCCCCTTTG	ATTCATATTG	GTAAAAAGAG	183840
GTTCAATCCA	AGTCTTAATT	TTATGACCAA	GGTCAACAAA	AATTTTACTT	AAATCTTTTTT	183900
CCGATTTTAT	ATAAATTTGC	GAAATTTCCA	ACACTTCACC	CAAAGAATAC	CTTAAATCAG	183960
CATTTTCAAG	AATCTTTGCC	GCAAAATTTA	GGGCAGCAAC	AACAGAAGAT	TCTTCTGTTG	184020
CAATTGGCAA	AGAATAGTAT	TTGCCATTTA	TTTTCAAATT	TTTTACAATT	CCAATAGGAA	184080
AAGATAAATA	TCCAATATAA	TTTTCTATCA	TATTAAAAAG	AAAATCTTCA	TTGGCATTAT	184140
TATAAAAAAA	ATCTTTATAA	GATAATTCCA	AAAAACTTTT	TATCTCTTGC	CTTTTTTCTA	184200
AAACGCTTTT	ATGTCTAAAA	TTTTTACTAA	GTTCCATAAA	ACTGCTTAAA	GACTCCAAGT	184260
TCATCAAGCA	AATAACTACT	TAAAAAATAC	TTATTATTTT	TAAACTCTAA	TAAACTTTTG	184320
CTTCCACTTA	AAAACATAGA	CATTTTTTAA	ATATGTTTCAT	AATCAGAAAA	AAGACCAAAT	184380
ACAGCATCTT	CTCCTGAATC	ATAAAAAGCC	CTAAGAACAA	CTGCTGCAAC	ACCTATAAGC	184440
CTGGCTCCAA	GGGCAATGCC	TTTAGCAATA	TCCATGCCCG	TCTCATATCC	ACCAGATGCA	184500
AAAATATTAG	CCTTTAGAGA	ATCATCAATA	CTAAGTAAAG	TAAAAACCGA	AGGTATACCC	184560
CAATCAGAAA	AACAAGATGC	AATGTTTAGA	TTATTACTCT	TCATGCCTTC	TACTAAAATC	184620
CAATTAGTTC	CACCACTCCC	TGCAAGATCA	ACATAAGAAG	CACCAAGGCT	GAACAATTCC	184680
TTAACGTCTT	TTGGCGAAAT	TCCAAAACCT	GTCTCTTTAA	CAATCAATGG	AACACTTAAA	184740
AAGTCTGACA	ATTTGGCTAT	TGACTCTCTT	ATTCCTTTAA	AATTTCTATC	TCCATCAACC	184800
TTCATCAATT	CTTGTCTTGC	ATTAAGATGA	ACAATAATTG	CATCAACTTC	TAATCTTTTA	184860
ATCATTTTCA	CTATTTTAGA	AATACCAAAT	TCAACAATCT	GAACAGCACC	AACATTGGCA	184920
AACAAAGGAA	TATTATGAGC	ATACCTTTTA	AGAGTAAAGT	CTCTTATGTA	CTCGGGATAC	184980
TTAAACAAAA	GCTTAAAAGA	ACCTAGCCCT	ATAGGAATTT	TTAAATAATT	TGCAATTCTA	185040

ACTAAAGATT	TATTAAAGTC	ATTCCCCTCT	TTACTGCCCC	CTGTCATGGA	AGAAATAAAA	185100
ACAGGCATAC	TAATATTGTA	TCCAAATATC	TCTTCTTTTA	TGTTTATCTC	GGAAAAATTA	185160
AAATCACTAA	GAGCATTGTG	TTTtagCTTA	ATAAACTTTA	AGAAATTACA	GCCACCTTTA	185220
ACATCGTTTT	TATTTAAACA	AATCTCAATA	TGCCTTTTTT	TATTTTCTAA	TATATTAGGC	185280
TCGATACCCA	TAAACTCGGT	ATCCATCATT	CCTTAGTTCT	TTTAAATAAA	ATCCTCTGGA	185340
TTCACCAGGA	ATTATTTTAT	TTTGAAAÀAA	ATCTTTATAT	TCTTCAAAAT	TTGCATTATT	185400
TCTATTTTTT	ATAAGACCTT	CAAGATCCCA	TAATTTAATA	ACATCAAAAG	CACTCTTTTC	185460
GATGGTAAGT	TCATAAATAA	TCATAATATT	GCCAGATCCA	TAAGAGCAAA	ACAATATCTT	185520
TTCCCCTGTA	ATATCTTTCT	TGGAAAATAC	TCTTTTTTAA	TAAAATGCTA	AAGATAGAAA	185580
AATTGAACCT	GTATACAAAT	TTCCCACTTC	CATAGCAGCT	TCAACTCCAT	CGTAAAAATC	185640
TATTGATTCT	AAATAAGCAT	TTCTAACAGA	TTCGTCATCG	CTATAATATT	TTTTCAAAAT	185700
ATAATGCATT	GAATCTATTG	GCATTTTAGC	AAAAGGAACA	TGCAAAACAA	ACCTATAATT	185760
AGAAAATAAA	TCTTTCATAC	TAAGTTGCTT	TTTGAAAGCA	AAATCTCTTA	AAGCATTTTC	185820
GTTTGcATTA	TTGTAACATT	CAACTGAATA	CTGACCTCGC	ACCTTAGCCT	CAACACTTCC	185880
AAAAGGCCTA	AAAAAATCGT	CAACATCATC	AGTATAAACT	CCAAATTCAG	ATAAATTGAT	185940
CGAAAGTAGC	TTTGGATTTT	TTTCAATCAA	AATTGCAGTT	GCGCCGGCTC	CTTGGGTAAT	186000
CTCAGCCGTA	GTAAGATTGC	TATAATGTGC	AATATCTGAA	GAAAAAACTA	TGCCGTATTTC	186060
AGAATTATTG	GAATGGCTTA	AAACACTTGC	CACAGTGTGC	AAAGACATAG	CAGCACCAGC	186120
ACACGCATGT	TGAACCTGGA	AAGTTAGAAA	ATTATTTCCC	AGACAAATAC	CAGATTGCTT	186180
TAAGGCTCCA	AAAACATAAG	AAGAAATGGC	CTTTGAATGA	TCAACGCCTG	TTTCAGTTCC	186240
ACCCAAAAGT	ATTCTAATTT	TGCTTAAATC	AAGATTATTG	TTGTCAAAAA	TAAGCTTAAC	186300
AGCCGAACCT	GCCATGGTTA	CACTATCCTC	ATTAGGACTG	GTAAACCTAA	AACCTTTTTG	186360
CAAGGTTGCA	TCTATTGCTC	TATTGATTTT	TTTAAAAAAA	ACTTCATTAG	AAAAATATAA	186420
AGGATTTTCC	AAAAGAACAG	AAAAATCTAA	ATAATTTAAA	GGTAAAAAAA	TTCTAATATC	186480
ACTAATACCT	ATTCTCATAT	ACTCCTCAAT	GAATTAATGG	CCTTAAGTAT	TATATTATAA	186540
TTTACAAAAA	TTAGCAAAAT	CTTATATAAT	AAAACCTAAA	AATGGAAGTT	TATGAAAATA	186600
GCCGTGCTTT	TATCTGGAGG	AGTCGACAGT	TCTGTTGCCC	TTTATAGAAT	TATAAACAAA	186660
GGATATTCAA	ATATAAAATG	CTACTATTTA	AAAATCTGGG	TTGAAGATGA	ACTGTCTTAT	186720
ATTGGAAACT	GCCCTTGGCA	AGAAGATTTA	AATTATGTTG	AAGCTATATG	CAACAAATTT	186780

AATGTACCGT	ATGAAATAAT	AAACTTTCAA	AAAGAATATT	ATAACAAAGT	AGTAAGCTAT	186840
ACTATTGAAG	AACTTAAAAA	TGGCAATACC	CCAAGTCCAG	ATATTTTTTTG	CAATCAAAGG	186900
ATAAAGTTTG	GAGCATTTTT	TGAGAAAATC	AATAGCCAAT	ATGATTGGT	TGTAACGGGA	186960
CATTACGCTA	AAATACAAAT	AAAAGAAAGT	AAATTTTTAT	TAAAACAGGC	AAAAGATAAA	187020
ATTAAAGACC	AAAGCTACTT	TTTATCTCAT	CTCTCTCAAA	AACAAATGTC	AAAACATATAC	187080
TTTCCCTTAG	GCACATTACT	TAAAAGCGAA	GTAAGACAAA	TAGCTAAAAA	CATAAATTTA	187140
CCCAACAAAG	ATAGAAAAGA	TAGTCAGGGT	ATTTGCTTTT	TAGGAAAAAT	TAAATATAAC	187200
GAATTTATCA	AATACCATCT	TGGAGAGAAA	AAGGGAAATA	TAATTGAAAA	AGAAACGGGA	187260
AAAATAATAG	GAATTCACAA	CGGATATTGG	TTTTTTTACAG	TTGGACAAAG	AAGAGGAATA	187320
AAACTTAGCA	ACGGGCCATG	GTTTCGTCATA	GAAAAAGATC	TGGAAAAAAA	TATTATATAC	187380
ATATCCCATA	ACGAGAATTA	TTTAAAACAA	GCAAAACGCA	AATTTTGTAGT	TCACGAAATA	187440
CATTGGATAA	ACGACACACC	TACGAACTTT	GAAAATTTCA	AAATTAAAAT	AAGACATGGC	187500
GAAAAGAAAT	ACTCATGCAA	ATTAAAACTT	ATTACAAATA	ACTTAATGGA	AATTTCTTTA	187560
AACAAAAAAG	ATCAAGGAAT	CTCCCCAGGA	CAATTTGCAA	TTTTTTTATAA	AAACACAGAA	187620
TGCCTGGGGG	GTGCTAAAAT	TTTTAAAATC	ATAGAATAAT	AATCCGCCCA	AAAAGTTAGA	187680
GAAGATTTTT	CAATCTTCTA	CTTACTTTTC	GATCTTAAAA	TAATCAACAG	ATCTTTTAA	187740
ATCTTTTACA	CTCTCTAACA	TCTTTTCAGA	CATTGCAGAA	AGCTCTTCAC	TGCTTGAGGC	187800
TGTAGTTTGG	ACTAACTGAC	TAACCTGCTC	TATTGCATTT	TTAAATTGCT	CTATTTGAAC	187860
ACTTTGCTTA	TAACTTTCAT	TAGAAATATT	TTTTACAAGT	CTGGCTGTTT	GTTCCATACC	187920
AGGAACTATT	TGTTCAAAAT	TTTCCCCAGC	ACGACTTGCA	ACAGTTAAAC	TTCTGTTTGC	187980
AATATCAATA	ATCTCTCTTG	CTGATTCTTT	GCTTTGATCT	GCAAGCTTTC	TAACCTCAGC	188040
AGCTACCACT	TCAAATCCCT	TGCCCTTTTC	TCCCACTCGT	GCAGCTTCAA	TCGAGGCATT	188100
TAAAGCAAGC	AAATTGGTTT	GCCTTGTTAT	CTCATCAATA	ATTCCAATTT	TTTCAGTAAT	188160
TACAGTCATT	GCCTCAATAG	CCTTAACAAC	AGATTTATGC	CCCTCTTTAG	TCCTTTCATT	188220
AGTATTAACA	GCAATTTTTT	CAGTAGTAGC	TGCATTTTCA	GTATTCTCAG	AAACACCTTG	188280
TGAAATTTGC	TCAATATTTG	CTGTCATTTG	CTCTAAAGTA	GAAGCCTGCT	CAACAGCGCC	188340
AGAACTTAAA	TTCTGGCTTG	CATTTGCTAT	TTGAATTGCA	TTTTCATAAA	GATAATCTAG	188400
ATTTTCAATA	ACTCCTTTTG	CAACTGAAGA	AAAATTGGTT	CTCAACTGCT	CAAGCCCTTC	188460
GTACAAACTG	TAAAGCTCTA	CAGTATCCCA	TTTGCCAAAA	TTAATATCAG	CAGTAAAATT	188520
ACCAGAAGCA	AGTCTCTCAG	AATATTCCAG	TATCTTATTC	AAAGAAGAGC	TTAACTTTTT	188580

CACAAGATAA	AGAGTTGCAA	TAGCAAGCAT	AAGTAATGTA	AATACAAAAC	TAATTGCTAA	188640
GATTATAGTT	GTAGCTCGTG	ACATGTAATA	AAAATCGTCC	TCTGAAGTTC	TCATTAAAAG	188700
AATAAATTTA	TTATTAGACA	AGTTTAATAA	CACCTTTTGA	CTAATTCCCA	CATATTTCTT	188760
ATTGCTTTTA	GGATCATAAT	AATAAACAGT	TGAAATTTCT	TTATTCTTTT	GCAACAAATC	188820
TTCAGATGTT	TTCTTAATAA	TATTGGAATA	AGAAGCACTA	ATATCAGTCA	AAATATCACC	188880
TGGCAATACT	ACATGATGAA	CCAACAATCT	ACCCGTAGTA	TCATAAGCTA	GCGCACGACC	188940
GCTAGAGAGT	ATTCCAAAAT	TAATCCCTCT	AAAAGACCTA	TATATATAAT	CCATTGAATA	189000
AAGAAACATA	AAATAACCAA	AAATGTTATC	TGTTTCAAAA	TCTCTTAATG	GCATACCTAT	189060
TAATATGTAA	GGAATCTGAC	CTTTTTTATT	TTTTATCTTG	GAAATATCTT	TAAGTAAAGA	189120
CTCCTCAATA	GACCCGGGAT	CTGCCAACAT	GACAAAGGAA	TTGTAAACAA	TACTATTAGA	189180
CTCCTTAAGT	TTTGTA AAAAT	ATTCTCTATC	CCCAATAGAT	TTGCCAAAAT	CACTATTATC	189240
CTTAACCGCC	GTAGTAAACA	CTATTTTACC	TTCTTTGTCT	GTAACCATCA	TATTACTACC	189300
GGTTTGAATA	AGAATTTTCA	ATTGATCTGA	AATAAAGCTC	AACCTTTTGG	ATTTAATTTT	189360
ACTAGCCTCA	TTAAATTTTT	CAGCAGAGTT	GAAATACATG	GAAC TAACCC	TAAC TTTCTC	189420
TTCCATTGAA	GAGATGTAAA	GATTAAAAGA	ACTTTTAATG	CTTTCAATAA	GATTTATCAT	189480
AAGATTAAAC	TGTTGATCCA	CCAATTTACT	ATTAATAAGC	ATTCCAAAAG	CAAAAAACAA	189540
AATTGATATA	AAGAATGCTA	TCAGAATAAG	AACAAGTAGC	AACATCCTAG	CTTTAAGCTT	189600
CATACTAACC	ACCTCTTTTA	CAAAAAATAA	ATTCTAAAAC	TCTGAAAAAT	CATCATCAAA	189660
ATTTAAATCC	TTATCAGCAA	TATCGATAGC	TTTTTTTAGGA	TCAACTCGCT	TATTAATAGT	189720
TCTTACAGAA	GATTCGCTCT	CAATATCTAA	AGAATAATTA	TTATGCCCAC	TGGCATTGTA	189780
AGTAGAAATT	CCATTGCTTT	TCAAATTTTG	ATTTTCATCT	TTAAAAGAAT	TTTCAGGACA	189840
ATCTATTAAC	CTGAAATCAT	AATCATCATT	TTCTGGATTT	TCAATTTTAG	AATCTTTAAT	189900
TTTGAAAAAT	AATACAGATT	TTCTAAGTTC	CTTAGACTTT	TCTAACATTT	TATCGGACAT	189960
ACTAGAAAAG	TGCTCACTGC	TTGAAGCTGA	AGATTGAACA	ACTTCTCCAA	CCTGATCTAA	190020
AGCCATTTTA	AATTGAGCAA	TCTGATCGCT	TTGCTTAGAG	CTACCTTCTG	AAATCTTCTT	190080
AACAAGATTA	GCCGTTTCTT	CAATTTTCGGG	TAGCATTTCT	TTAAAGATCA	CTCCCGCTTC	190140
AGTTGCTACC	TTAGAGTTAT	CTTCAACTAA	CTCTCCAATC	TCAAGAGCAG	AAATTTTACT	190200
CAAATCAGCC	AACTTTCTAA	TCTCACTGGC	CACAACAGCA	AATCCCTTTC	CCTCATCTCC	190260
TGCTCTTGCA	GCTTCAATAG	CCGCATTCAA	AGCAAGTAAA	TTGGTTTTTC	TAGCTATCTC	190320

TTCAATAACA	CTAACTTTCT	CCACAATGTC	TTGCATAGCA	ATAACAGATT	CTTCAACGGC	190380
CCTACCACCT	ATCTGAGAAT	TTTCATTCGT	CTTTAAAGCT	ATTTGTTCTG	TTTCATAAGA	190440
ATTATTGGCG	CTCATGTTGA	CACCTGAGGC	TATTTGCTCA	ACATTAGCTG	ACATTTCTTC	190500
AAGAGCAGAT	GCCTGTTGCA	ATGCACTAGA	GCTTAAATTT	TGACTTGAAC	TGGCAACTTC	190560
TAAACTTGCC	TTATTTACAT	AGCTAATATT	TCTCAAAACA	CTAGAAATTG	CTACAGAAAT	190620
AGCTTTTTTC	ATTTTAACAA	CCTGAAGACT	TAACATGCCA	AGTTCATCAA	GAGTATTTTC	190680
ATCATCATCA	AGAGCATAAT	CTTTATCTAA	ATTGCCCTTA	ACCATATCTT	GAAGTAGAAC	190740
TCTAATTGCG	TTTAAACGAA	AACTAATAAT	CCTGTCTATT	CTAATTGAAA	GAACAATACT	190800
TAATGCTATA	ATGCCTAAGA	CTGAATATAA	AATATACTGA	AATCTTAGAC	TAGATATTAC	190860
TCCGTAAATA	TCCTTATAAG	GAAGCCTAGC	AATAAGTACT	CCACTCTTTT	CTCCCAATTT	190920
ACTACTTATG	GGCAACATTG	CATAATAACA	ATCTTCTCCC	ATTTCCGACA	AAAGTATTCT	190980
ATCAATAGTG	TAAACCGACA	CTTCACTGGC	AATGTTTGAT	GGAAAAGGGG	GCTTAGAGAA	191040
AACATCTTTA	AGAACATTCA	AAAATTTAGA	ACTAACCCTG	CTGGTTTCAT	TATATTCTTC	191100
AAAAGGATTA	ACTGCTATAT	TGTTGGGATC	CACATAAATA	AAATTGCCTC	TTTTATAAAA	191160
ACCGAATCTA	AATCTATCAA	AACTATCTGC	CACAATATCA	TTAAGCAAAT	ATCCGGCCAA	191220
ATACCCACAA	ACAAGTTTAT	CTTCTGGGGA	ATATACAGGT	ACAATTATTG	CAAAAGCCTT	191280
TTTTTTCGCTT	TGTTTAGACC	TAATAGCAAC	TTCTGCGGAT	ATTCCTTCAG	AAAGATTTGA	191340
ATACCAACCT	ATAAATTTTA	ATTGGTTTTG	CCTATAATCC	TCAACAGCTT	TTTTAAATA	191400
ATTGGTATTA	GCCTCAGAAT	GACCAAAATC	CATATTATTC	TCATGTCTTG	TGCTAACAAT	191460
TACTCTCCCT	TCAAAATCAA	AAAAAGCGAA	TTCTTCAAAA	AGGGTATCAT	TTTTAAGATT	191520
GGCCATAAAA	TTGTATAAGT	ATTGTCGATA	TTTTTTGCTC	ACCTTTACAG	AGTCAATAAC	191580
AAATTTTGGA	TTTTTTCTTA	AATCTATCAA	TTCCGACTCA	GAGAAATCTT	TTCTCTATT	191640
CTCAGACATT	GCAAATTCTG	ATATGGTTTC	AAGTGCCAAA	TTAGAAGCTG	CACCATTGAT	191700
TATGACATGC	AGGGTGTCTA	AAAAAGATTG	CAAAGAAAAA	GCTGCTCTTC	TTACTTGCGC	191760
CCTTGTAAGC	TGCTTATAAT	AATCTTCTAA	ATAACCGCAT	AAAACAAAAT	TAAAAATCGT	191820
GGAAAAAGT	AGCAGTATAA	AAATTAAAAA	CAATAATAAA	AATCCAACAA	ACCTGTATTT	191880
AAGCTTCAAT	AACATAATAA	ACTACCTCAC	AAATCACCTA	CTTATTTAAT	CAAATAAACT	191940
CAAGACCAA	GGGTATCAAA	AAAAAATTTA	CAGCAAAAAA	TCAAAAACCTC	TCAAAAAAAT	192000
AAAAGAATTT	TCACAACATG	AAAAACAAAT	TATAAACTAT	TATTATTAAC	TGCTAATGCT	192060
TTTTATAGGT	TATTCCTAAG	AACTTAGGCG	CATAATTTTT	TTTTGAAAAA	AAGATAAGAC	192120

TTAGAATAAT	AATTAAATAT	GGGGTAATCA	CTAACATTTT	GGGAGGCATT	ATTAAAGAAA	192180
AAAAGGGTAA	TTGAGCCAAA	ACAATTGCCA	AAGTTTTTCAC	AAATGAAAAC	AAAAAACTGC	192240
CTATTAAAAC	TCCCAAAGGC	GTCCATTTTC	CAAAAATCAA	CATTACAATA	GCAATAAAAC	192300
CTTGTCACC	TGTAACCCCT	TGCACATAAC	TTGATGCAAC	CACCGTTGTA	AGAACAGCAC	192360
CTGAAACCCC	TGCTAAAAAA	CCACTCAAAA	GAACGCAAAA	AAATCTAATT	TTATTTACAC	192420
TAACTCCAAC	AGACTCTAAT	ACCTCTGGAT	TTTCACCACT	TGCATTAATT	CTAAGCCCAA	192480
TTTGTAGTGA	TTTGAAAACA	ATATGAAACA	AAACCACACT	TAGGATTGCA	ATGTATACAG	192540
AATATCTTTT	GCCAAAAATT	TGAAAAATAA	AAGATGTTTT	GTTTAAAATT	CCATCAAAAA	192600
GTATCGGCAA	CTTTATTTCT	ATAGGCGGAG	TTGAAATAGA	AGAAAAAATC	AAAGTGCTTA	192660
TAAAAACAGC	AATAGCGGGT	CCTAAAAAAT	TAAGTGCCAT	TCCGGTTATA	ATTTGATCTG	192720
ATTTTAAAAA	AATTGTAAAA	ACAGCGTGCA	AAATAGCAAG	AACAAGCCCT	GCTAGCCCAC	192780
CAGCAAAAAT	TGAAAACAAT	GGATCATTTG	TAAAATATGC	AACTGTAGCT	CCTGAAAATG	192840
CTCCTATTGT	CATTATTCCT	TCAAGTCCAA	TATTAATAAT	TCCACTTTTC	TCGCTTATAA	192900
GACCCCCAAG	ACCAGCTAAA	ATTAAGGTTT	GAGAATTTAT	TAAAGTTTCA	CTAATCAAGA	192960
ATATTATTAT	ATTTGACACG	CTTAACACCT	TTTAAACAA	TTTTATTTAA	AAAATAGCTA	193020
GCAGAAATTA	CAAGAACAAT	TATTCCCATC	ATCAAAGATA	CAATTGAAGA	TGGAAGGCCC	193080
ATTAACTTTT	GAACCTCTACT	GCTTCCATAA	AGCAATATAG	AAAAAAGAAT	GCTAGAAAAT	193140
ATTATGCCAA	TTGGCGAATT	GTTTCCCAT	AGAGAAGCAG	CTATCCCATT	AAAACCAATT	193200
CCTTGCATAT	AAGAAAGCTT	AAATATAGCT	TTATTAACAC	CCATAAGTTG	AATAGCACCA	193260
GCAAGACCTG	CAACAGCTGC	TGAGAGAAAC	ATTGAAAAAA	TTAGCACAGC	TTTTACATTA	193320
ATACCCATAC	ATCTTGAAGC	TTCAATATTA	CTTCCTGTGG	CATTTATTTT	AAATCCAATA	193380
ATAGTTTTAT	TAAGTAAAAA	CCATATTAAA	ATAGCAAAAA	TTATACCTAA	AATTATTCCA	193440
AAATGAAGAG	GTGCTTTTAA	AAGCTCATTA	ACAAAAGGAT	GAGAAGATCT	ATAAGCAAGA	193500
CCTTCTGGTG	AGAGCTTCCA	AGAAGCTAAA	AAATCAATAT	ATGCGCTTTC	TTTAATGGGT	193560
TTTGAAAAAT	CACTATTATC	TCTTTTAATA	AAACTAAAAT	CTAAAATTAT	ATTATTTTAA	193620
TGAAATAATA	TCCAATTAAA	CATTATTCCT	GAAATCACTT	CGCTAATATT	GAATTTGGCT	193680
TTTAAATATC	CGATTAAAA	TCCTAAACTG	CCTGAAGCTA	AAAAAGTAAT	AATAAAAAATA	193740
GTAATTACAT	GTAAAAATTGG	AGGCAAATCA	AGTAAACTG	ATGCTATTAA	AGCAACAATA	193800
GATCCTAGTA	TAAACTGGCC	TTCAACCCCA	ATATTAAAAA	GACCCGCTTT	TAAAGAAATA	193860

CCAATAGAAA	GACCTGTAAA	TTCAAAGGA	GCTGAATAAC	TTAAAACATA	CTTAAATGT	193920
TTGGGAGAAG	AAAAAATAAT	TTCTAATATT	ATAAAATACA	TTCTAAAAGG	AGAATGACCA	193980
AGCCCCATCA	CCACTAGCCC	AACAATTAAA	AATCCAACAA	ATAGAGCAAA	TACACTAACA	194040
AATGCTGAAG	AATTTAAAAA	TTTCAAAATA	AATTTACTAA	ATACGTTTTT	ACTAATTGTC	194100
ATTTAGCTTA	AGCCTATCAT	CATTTTACCA	ATAACATCAA	TATCAAAATT	ATCCTCTAAA	194160
ATGCCCCACTA	TTCTTCCACC	ATGCATTACA	GCTATCCTGT	CACAAACATT	AACAAGCTCA	194220
TCAAGTTCAA	GAGAAACCAA	TAAAACAGAT	CTACCCGCAT	CTCTTTGCTC	TATTATTCTT	194280
TTGTAAATAT	TCTCAACAGC	TCCAACATCA	AGGCCTCTTG	TAGGCTGAAT	AGCCAAAAGA	194340
ATATCTGGCT	CTAAACTAAT	CTCAGGAGCA	ACAATAACTT	TTTGTGATT	ACCTCCGGAT	194400
AAATGCTTTA	CCTTGCTTAA	AATATCTCTT	GGTCTAATAT	CAAAATGACT	TACAAGTTGA	194460
TTGCTCAATT	TTCTTAAAAT	ATTAAGATCA	AACCCACAA	ATTGTTTTTT	AACTTATTG	194520
AATTGTCTTT	TAATAAAATT	GAAAAAATTG	AATTTTAAAT	CAAAATTACT	CTTCAAATGA	194580
ATTGTTTTTA	ATCTCAAATA	ATCAGGATTA	TCAAAGCTCT	TAAGTCCAAT	ATTTTGCATA	194640
ACATTGAATT	CTAAAAAAG	GCCGTGTTTT	TGCCTGTCCG	AAGAATATTG	CCAATTTTTT	194700
TATCTATTCT	TTGTTTAATA	GTAAACCTT	TTAAAGATTC	CAAATTCCCC	GAAGAATTTT	194760
TTTTCAAAAT	ATCGCCCTTA	AATATGCTTT	TCAAACCCAA	AATTGCATCA	ACTAAATCCT	194820
CCTGACCACT	CCCCTCAATA	CCTGATATTC	CAAGAATTTT	TCCATTTCTC	AGATCAAGAT	194880
TAACGTCTTT	AACTTTTAAA	ACTCCTCTCT	CATCTTTAAC	ACTTAAATTC	TTTATTTCAA	194940
GAATATTAAA	ATGATTTTCA	AATTTAATTT	TAGATGAGCG	AAGTGCAACT	TCTTTTCCTA	195000
TCATTAATTT	TGTAAGATCT	TTGTCATCAA	TATCAGCAAT	ATTAACAGTT	TTTACAACCT	195060
TCCCAAGACG	CATAATTGTA	CATTTCTTTG	CAATAGATCT	AATTTCTTTT	ATTTTATGGG	195120
TAATAAGTAT	TACAGTATGA	CCCTCTTGGG	CGAGTACCTT	TAAAATATTT	ATAAAATCAT	195180
CAACTTCACT	TGGAGCAAGC	ACTGCTGTAG	GTTTCATCAA	AATAATAATA	TCTGCATTTT	195240
GATAAAGAAC	TTTCAATATC	TCTATTTTTT	GTTCCATGCC	AACACTCAAG	TCTTCAACCC	195300
TTTTTTCTAA	ATCTATCTTT	AAACCATACT	TTTCCGAAAG	AGAACTTATC	TTTTTTCTAG	195360
CTTGTTTGTA	ATCAAGAAAA	CCAAATTTTG	AATTTTCATA	TCCCAAAATA	ATGTTTTGAA	195420
CAGCAGTAAA	TTGCGGAATT	AACATAAAGT	GTTGGAAAAC	CATTCCAATC	CCATTTGCGA	195480
TAGTTCGCTT	GAATCCTTAA	AGTTTATTTT	TTGACCTTTT	AAAATAATTC	GACCACTATT	195540
TACTTGATGA	ATCCCATAAA	TAGTTTTTCAT	TAAGGTAGTC	TTTCCAGCAC	CATTTTCTCC	195600
AAGAATAGCA	TGAACTTCGC	CTGCCTTAAA	TTTAATAGAA	ACATTATCAT	TGGCAACAAA	195660

ATCACCATAC	TTTTTTGTAA	TATTTTCTAA	TACTAGTACA	TCTTCTTTCA	TCAAGCTTTA	195720
ATCCTAATAT	AAATATCAAA	CATAAATGAC	TTACAATATT	CTATTCTAAT	GAAATATAAT	195780
AAAGCTTAAA	ATTTATATTT	TCAATAGCAT	TTTAAATTTT	AATAAACAAA	ATAAATTATC	195840
TATAAAAATT	AATTGCTAAT	AAATAAAAAT	GCTTAAAATA	TAAAAACCTT	AATAAAGAGA	195900
GCAAATTAAT	GTTAAAATAA	AAGAATAAAT	AAATTATTAC	AAAAGAGAGT	ATTATGAAAA	195960
TCAAAGCCTG	CATTTTTGAT	ATGGATGGAA	CACTGGTAAA	TAGTATAATG	GATATTGCAT	196020
TCTCAATGAA	TTCTGCTCTT	TCAAACCTTAG	GATACAGTAA	AATAGAACTA	AGCAAATTCA	196080
ATGCCCTTGT	TGGCAGAGGA	TTTAACAAGT	TTGTAATAGA	CACTCTAAAG	CTATTATcTC	196140
TTGAACATGA	TAATCCTAAT	TTACAAGAAA	AACCTTTACAA	AGAATTTGTT	AAAGAATACA	196200
ATAAAAACCT	TTCATTCCAA	ACAAAACCAT	ATGAAAATAT	AAAACCCCTT	TTAGAAACTA	196260
TGAATAAGCT	TAACATTCCA	ATTGGAATTT	TAAGCAATAA	GAACCACGAA	GAATTAATAA	196320
ATTTGGTGAA	AAATATTTTT	GGAAATATAT	TGTTTTTTGA	AATCAGGGGT	TATTCAAAAA	196380
ATTTTCCACC	AAAGCCAGAT	CCTGAAAATG	CCCTTGATAT	GATATTAGAA	TTAAATGCCC	196440
AAAAAGAAGA	AATTGCATAT	ATTGGAGACA	GCGATGTGGA	TATGCTAACC	GACTTAAACG	196500
CTGGATTTAT	GCCAAATAGG	GTTTCTTGGG	GATTTAGAAG	TGTTCAAGAA	TTAAAAGAAA	196560
GTGGAGCAAA	ACATATAATA	CACAATCCAC	TTGAACTATT	GGACCTAATA	AAATGAATAC	196620
AAAACCATAT	TTTCCTTATT	TATATCATT	TCTATTCAAT	CATGAAAGTA	TAAAAAGTTT	196680
ATCTGCCATA	GAAAAAGAAA	TTGAAATACT	CAACTATTTA	AAAGAAAACA	AAAAACAAT	196740
TGCTACATTT	ATCAAAAATG	ATTTTGAATC	AGAAATAAAA	GACCTAATTC	AATACGTCAA	196800
GGATAAAACA	GATATAATGA	TTACCCCAT	TGTTTTATCT	GGCATTGAAG	CTATTGATTT	196860
TAACATTGTA	AAGCCTCTTT	TTAGTAAAGA	ATTGACAAAA	AACGACTTAA	ATTTGATATT	196920
TAACCTTGTC	AAAGTCAACT	CATCTTTAAG	AAAAGAATTC	TTTTATAATT	TTAATACCAT	196980
AAGCAATGGA	TACATTACTT	TTTATATAAA	CAAACATTTT	GAAGGAAAAA	ACTCTTATAC	197040
AATATACTTA	ATACAAAAGG	AAAATAAAGC	ACTTTATTCA	TCAGACATCA	TAAAAAATTA	197100
TATAAAGATA	CTACTTCTCT	TAAAAGTATT	GGTAATTAAA	TACTGCTTTG	AAAAAGGAAT	197160
AGAGCTTACT	ACTAAAAACA	TTGAATCCAC	TTCAAAAGCA	ATAAGCAATG	ATACCGACTT	197220
TCTAGACGAA	AAGACAGCTA	AGCTTATAAT	TGAAAGCTTT	TTCAAATATG	AGACCTTACA	197280
AACAATGTCT	CCAATTTCAA	CATTAATTGC	CATTTTTTCA	GCCAGAGCAA	GAACCTCCAAA	197340
ATACAAAAC	AATCCGGTTA	AAGGTTTTAT	TGGGTATGAT	GAAAGTTGGT	TTTCAATAAA	197400

ACAGTCGGGC	TCTAGAGAAT	AGATTCAAG	AATAATTAAA	GAATTATCAG	ATATAGCCAA	197460
GGTAAATAAA	TGGTAAAAAA	ATTTTCAATT	TTCTTAAAAG	CAATAATAAT	TTTTTCAATA	197520
TTTGAACTTT	TAATCGAAGA	ACTCTCAATA	ATTCTTTTTT	TACCATACAA	AATACGATTT	197580
GCACTAATAT	TTCTTGGGTT	TCTATTTGAC	ACAATTTTTA	TTTTCATTTT	TTTATACAAA	197640
ATAACCAAGG	CCTACCTTTC	CCAAAGATTA	GAAATCTACG	TCAGAAACAA	TCTATTCTTC	197700
GATATAATCC	ACTGCCTTAT	TCCTTTAGCG	TTTTATAGCT	CATATCAGCT	TAAAAACATA	197760
ATTGTGCCCC	ATGAAACAAT	ATTAAATCCA	ATAATGCTAT	CACTTTTCAA	GTTAAGATTT	197820
TTAAGACTTC	TTAGGTTTAA	TGACCTAATA	ATAGAAATAT	ATTACAATTC	AAAAGAAAAG	197880
AACCTAATAC	TAATAGCATT	TGCTAGGACA	TTTTCAATGA	GCTTATTAAT	ACCATTTACA	197940
TTTTTTTATAA	TAATATCAAG	CTCAAAAATT	GTAAATTCAA	TACCAGAAAA	ACAAGAATTT	198000
AATATCATT	AAAATATATC	AATAATAAAT	GAAAAAGCTT	ACATTAAAGA	AAAATATCCC	198060
TTCATCTTAA	TAATCAAGGA	AAAAGATGAC	ATAATATACT	CAAAATCAGA	CGAAATATTT	198120
GTTTACTACA	GTCCAGTGA	ATATAGAGTA	ATAGAAATGG	AGAAAACAAA	ATTTTATATA	198180
GATAAATATT	TGCAAAGAAA	AAGCGATTCT	ATTCTTGGAA	TTTTTCTATT	TACATTGTTT	198240
GCATCATTTA	CTATTTTTTT	AATGAATTTT	TATAAATTTT	TTAAAGCAAG	CTTTTTTAAAT	198300
CCTATTATTT	TAATGACAAA	AATTTTACAA	GACCCATTAG	AATATCGAAA	AATTCAAATT	198360
CCTTTTACTT	TAAGCGAAGA	AAAAGTATAT	GAAGTTGCAA	AATCATTTAA	CAATCTCTTG	198420
CTAAAAGAAA	AACTAAACTC	AAAGCGAAAA	AGCAAAATAC	CTTTAGAAAT	TGAAAAAGTA	198480
AAAAAAATAA	TTAATAAAAA	CCAGGAAATA	AAATGAAAAT	TCAAATAATT	ATAATGCTGC	198540
TTGCATTGTT	AGATTTTCCA	CTTAATGCCA	GACTTTTGGG	CATTTCAATT	GAAAAAAGAG	198600
CAGATGAAGA	AATAAAAAAA	TATTCGTCTT	ATAATTTAAT	TTTAGAAAAA	GAATACTATA	198660
CCAATTTTCC	AACAAGCGAA	ATAGAAAAAA	ATATTTATAA	ACTAACAGAA	CATTTTGTA	198720
AAAGCATAAT	GCTCAATAAA	ACTAACTACA	GCTTATTAAA	TTCAAACACT	AAAGAAGCAA	198780
ATAAATATCT	AATTCAAAGC	GAAGTCATTG	ATAAAAAATT	TTTAAATAT	AAAATATTTA	198840
AAATCAAAAA	TATAAATGGA	ATTTTTAAAA	GCCATTCACT	AATATATACA	AAAAAAGGAT	198900
TTTACAAATT	AGAAGTTTAC	ATAGAAAATA	ATGCAGAACC	TCTAAAAATA	TTTAACCTTA	198960
ACATTACTTA	TTTTTTTAAAG	AATTTAGATA	AAATAAGTAA	TGAAATGATT	TTTTTCCCAA	199020
GGGAATGAAA	ATAATAAAAT	TAAAGCTTGA	ACTTTTTTTA	TAAATAATTT	ATTTAACAAA	199080
TACAGACATT	ACTCTTTGAA	GAACCTTTGC	CCTATCTAAT	GGTTTAACAA	TAAATGTTTT	199140
TGCTCCTTTT	ATTAAGCAAT	CCTTAACTAA	TTGTTTCCTG	CCTAAAGCAG	ATATCATTAT	199200

CACTCTAGCA	TTTTTATCAA	ATTCCATAAT	ATTAGAAAGA	CAAGTTATTC	CATCCATTTT	199260
GGGCATAGTA	ATATCAAGAG	TGACAATATC	AATATTAGGA	TAATGATTCT	TGTATTTTAT	199320
CACAGCCTCT	TCTCCATCAG	CTGCCGTATC	AATAATATTA	AAGCCCTCTG	ATGTAAAAAT	199380
TTGAGTAAGC	TGCTTTACGG	TAAAAACAGA	GTCATCAACA	ATTAAAAACAT	TAAAAGGAAT	199440
GCCTGTATCA	TAATTGATTC	CTCTAGGCTT	AGATGAAGAA	TCTGCAGCAA	TTGTAGTCTT	199500
TTGAATCATA	TTAACCTCTC	TCTTCTAATA	AAAAGAATTT	TTTTCATATC	AAACCCTCTC	199560
TCTTATTGCA	ATATTAACCT	CTATAATTTT	ACCATCAGGC	AAAGAAAAAG	GAACAATTAA	199620
AGCCTCAGAA	CCTTTATTAC	TTATTTTCAT	ATTTTCTCCA	TAAATAAAAAG	CTGGGGGGGT	199680
TATATCAAAT	ACAAAACCCCT	TGGCATGCAA	AGTGGTAACA	AAATTTCCAG	CAATAATATT	199740
GCCAACCTCA	GTTAGAGTTG	CAGCAACCAT	CTCTTTTGTT	TCTTCATCGT	CAAAATCATC	199800
ATACTCTTCA	AAATTTAATT	TAGAAGCAAC	AAAAGAGCT	GTTTCTATGT	CCATATCAAT	199860
AATTATACTG	CCCTCAACAG	ACCCAGCAAG	CCCTACTATT	ACAGAAACAC	CTTTTATCTT	199920
TTGATTTATC	GACTTAAGCC	CGGGCTTACC	CATTTCTATA	TTCTCAACAA	GCAACATATC	199980
TCTTAAAACC	GAAGAAGCAG	CATCCAAAAA	TGGCTCTATA	TAATCTATTC	TCATTAATTT	200040
CTCCTTTAGA	CTTTCCTGTA	CAAGTTAAAA	TATTTTGTGG	ATTTCTCTTT	TATAAAAACA	200100
TCATTATTTT	TAAGCTCTTC	GTTATCTCCC	AAAACCAAAA	GAGCCCCTTT	AATAGCTTTG	200160
GAAGCAATGA	TATTTAAAAT	TAAAATCTGA	TCTTTACTAT	CCAAGAAACA	TAAAACATCT	200220
TTTAAAAAAA	CCATTCCCAA	ATTATCAGGC	AAATCTGAAA	AAAGAGCATC	GGAATATTCA	200280
AACAAAACAT	TACTTAGGAT	TTCTGACTTA	AATTTATAAA	CTCCGGGACT	TTGTTCAAAA	200340
GAATTCCTAC	TATAAATTTT	ACTAATGCCA	ATTTCTGACT	CTGAAAACAC	CAACCTAGAA	200400
GTTTCAACAA	CTTTTGATAA	ATCATTATCA	ATGGCCGTCA	ACTTAAAAGG	TTTACATAA	200460
TATTCAGACA	AAGCATTGGC	TAAAGCCATA	GTCTCCTTTC	CACTACCACA	ACCAATTTCC	200520
AATACATTGA	AAATAGAATT	TAAGTTATTC	ATAAACTTA	AGCGACTCTC	AACAATTTCA	200580
TTTTTTAAAT	CTTCCAAACA	ATCAGCTCCC	CACAAATTTT	CCGATGATTT	TGAATAAAAT	200640
TCATTTAAAA	AACTATCGCA	AGGCAAGTAA	TCAGTATCAA	CCATATTAAA	TTTTACTCCA	200700
ACTTTTTTCCA	AAAACACATC	ATTTACCAAA	GAAGCATTAA	ACGAGTATTT	CAAAAGATTT	200760
TTTTTTAATAT	TTTCCAAATT	AAAAGCTGCA	GTATTGTTTA	GAGTTGAATT	TTCATTGTTA	200820
CCATCATTTT	TTGAAGAAAC	CTTATTGGAA	AAATCATTCT	TAGAATTTTC	TAGTATATCT	200880
AAATTATCAC	AATCGCTTAC	AAAGTCGCTT	TTTTCAACAA	AATTTTGACC	AGGTTTTAAT	200940

AACTTTTCTT	CTTCACCATA	TTTAAAAATT	TTAAAAACAT	TAAGAAGTAT	TTAAAGCTTT	201000
TCGTTATAAT	CTACAACGCC	TTTTATATAG	TTTATTAAAG	AATCCTGAGA	TAAAACTGGA	201060
TGTGGATCTT	GAATAAGGCT	AGAATCTATT	GAAAAAACAT	TATTAATTTT	ATCAACAATT	201120
ACCCCTATAA	GAAGGTCTTC	GTTTTTTAAA	ACCATAATAT	CTTCAATATC	TTTTTTATTA	201180
AATTCTAAAT	TAAACATTAT	TCTAAGATCT	ATAATAGGAA	TTATTTTACC	CCGTAAATTA	201240
TCAAGCCCAG	CAACATACTT	TTTGGCATT	GGAACATAAG	TAAAATTACT	AGATTTTCTA	201300
ATTTCTTTAA	CCTGCATAAT	GTCTACTAAA	TAATGATCCG	ACCCAAGCTC	AAAAGAAACA	201360
ACTTTAAAAAT	CAAAATTGGT	CAATTTAGAA	TTAGAATTTT	TATCATCTAA	AATTTTGGGT	201420
CCAAAATAAA	TTTCTTTTAT	CTGCACAAGA	GATCACTCCT	TAGTATCCTT	TTGTAAATCA	201480
AAAAGTTTAA	AAACATCAAT	TATCAATACA	ACCTTACCAT	TGCCAAGCGT	AGTAGCCCCA	201540
ACTATACCCG	CGCTTGATGA	AAATTTATCC	TTAATAGGCT	TTACTACAAA	ATCTTCCTCA	201600
CCAAGAATAG	AGTCTACAAC	AATTGCTATC	TTCATGTTGC	TAGTATTAAAC	AACTATTAAA	201660
AATTTTCTA	TTAATGAATC	ATCCCTTGTT	ATGTTAAAAA	GTTTATCAAG	CCTGAGAACA	201720
GAAATGACTT	CATCTCTTAA	ATTATAAACT	TCATGATAAT	TTTCAAGCAA	TTTTATATCA	201780
TGTTCAAGTTA	TTCTATGAGT	TTCAAGAACA	TTATTTAAAG	GAATAACATA	AGTCTCAGAC	201840
CCCGACTTTA	CTAAAAGACC	TTGTATAATC	ACTAACGTCA	ATGGTAGTTT	AATTTTAAAA	201900
ATTGTTCCAA	GACCAATTTT	TGATTCCACC	AAAATAGTTC	CATTAAGCTT	TTCAATGCTT	201960
TTTTTCACAA	CGTCAAGACC	AACTCCTCTA	CCTGAAAGGT	CTGTCACTTG	AACTGCTGTT	202020
GAAAACCCAG	GAGCAAAAAT	TAAGTTAATA	AGTTCAAAAT	CAGAGTAAAT	TGCATCTTCT	202080
TTTATTGTTT	CCTTTTCAAT	TAATTTGCGC	CTAATGACCT	TTGGATCTAT	ACCAATCCCA	202140
TCATCTTCAA	TCTCAATTGA	TATTACATTA	CCTTCATTCT	TGGCACGCAA	AATTATAGTA	202200
CCTGCTTTGC	TCTTTCCCCT	TTTAACTCTC	TCTTCAACTG	TTTCAAGGCC	ATGATCCATT	202260
GAATTTCTAA	CACAATGCAT	CAAAGGATCT	ACAAGGTCAT	CTATAACAGA	CTTATCAAGC	202320
TCAGTTTCTT	CCCCTTCCAT	TTTAAGATTC	ACAATCTTAT	TTAATTTCTT	TGAAAGATCT	202380
CTTACGACTC	TTGTAAACCT	TGAAAATATA	TTAGATATTG	GTAACATTCT	GGTTTTTAAA	202440
ACACTCTCAT	GCAAATCTGT	AATTATTCTT	GACAGCCGCC	CAGAGGTCAT	TTTAAAATTT	202500
TGAAGAAGTC	TGAAAAAAGA	ATTTCTCAAT	TCAGATATAT	CCTTAAGAGC	CTTTTGCATT	202560
TTAAAACTCA	TCAAAGAATT	AATATGTGAT	TCGATCTCAT	CTTCTAATGT	TAAGCCTGCA	202620
TCTTTGAAAA	CTATCTTTAA	ATCAATTAAA	AAGTTTCTCT	GAAAACTTTC	TTGATAATCA	202680
TAAAAATAAT	TAAAATTATA	AAATAATGTA	ATCATTTCTG	AATTTATTTG	ATTATAAGAT	202740

GATTTACTTA	TTACAGCCTC	ACTGACAAGA	TTTAATATGT	AATCTATTTT	TTTGCTATCT	202800
ATTCTAATTA	AATTAACACT	AATTGGACTA	TTTTTCTTAA	TATTTTATT	TTCCTTAAAA	202860
GGTGCTTCAT	CATCTTCTTT	TAGCCTTACG	CTCTTTAAAG	ATTCTAAATT	AACATTTTTG	202920
ATTTCAAAAT	GACTAACAAC	ATCTGGTAAA	TTAATCTTTT	TAGCAATACT	TTCTTCACTG	202980
GTATTTGATA	TTAAGTAATA	TATTACAAAA	TCAAAAAACT	TATCTGCCAA	TAATTCGCTA	203040
GAATCTGGGA	TAGACTTGAA	AATTTTACCA	AGACTTTTTA	ATGCTTGAAG	CATTTGAAGC	203100
CCACTAATAG	TAGCCATAGG	ATTGTCTTTT	ACAAAATCCA	ATCTAACTTT	AAATAACTTT	203160
TGATTTTCAA	CCTCAAGTAA	TAAATCAGAA	ATCTCATCCT	CTGTAAAATC	AAAATTATCA	203220
TTTAAACAA	AATTGGAATC	TAAATCAACA	TCTTTAATCT	CTTCATCTGC	AAGCTTTTTT	203280
AATTCCTCTT	TTACGTTAAA	TTCATCAACC	AAATAGCTTT	CAATTAAATT	TAAAGAATCT	203340
AAACTTTTTT	TCACGCCTTC	TATATCTGAA	TATATCAGAT	AATAATCAAC	CCTTTTAAAA	203400
AATTTATCCT	CTATGATTTG	CTCATATTTA	GGAATTGTAT	GAAGTACAGA	TCCTAAATTT	203460
TTTAAATAT	TAAATATTTT	TAGCCCACTA	TTTTCAACTT	CAGAATTGCT	ATTTGAATTA	203520
AAAACAACAC	TGATCCTTAA	AACCTTTTGT	CCAATTCCTA	GCCCTTCTCT	TATCTCCTCA	203580
AGATCTGATT	CTGAAAGACA	AAAATTGTTT	TTAATTGAAT	TTCCATCAAA	TCTCTTAATA	203640
AAAGCTGAT	CATCAATTAC	TAAAAATTGC	TTTAATTTGC	TTTTAAGATC	ACTTATGTCA	203700
TTTAAATAAA	CCTTGCCATC	AATACGAAGC	GCAAGCATTT	CCTTGATAAC	ATCTAATGAA	203760
CTTAAAAGCA	GATCAACAAG	ATCATTATTT	ATATTTACCT	TACCATCTCT	AATAGCATCA	203820
AAAACATCTT	CGACAATATG	GGTAAAATCA	GATAACTCCA	TCATATCAAG	AGAAGCAGAG	203880
CTTCCTTTTA	AAGTATGAGC	TGCCCTGAAT	ATTTTCATCA	TAGTATCAGA	ATTATTAGGA	203940
TCATCCTCTA	ATGACATAAT	ATTCTCTTCA	AGGATATCTA	CAAGATTTTG	AGCTTCTTCA	204000
AAAAAACTC	CTAAAAGCTC	TTCATTTTCC	AAATCTAATA	TTTCCATATA	CTATTTTCTA	204060
TTATTTTAAT	ATGTAAAGCA	AACCTTTTAG	GTAGCTTTAC	ATATTAATTT	AAACCTAATT	204120
TTTCGGAGAT	GATTCTTCAG	GTTTTTCACT	CTCAATCTTT	TCTACAAAGT	TTTGGAAGA	204180
GCCTTCAGGC	ATAGAAATTT	TTTCTCTAAG	CTTTAAAACT	CTTTTAAAAG	TTTCGTGTGC	204240
CTTTAATTTA	CGAAGGGATT	CAGTTCCGCT	AGTCTCATAA	ACTTTAAATA	CAGACTCACT	204300
GTCAATATCA	GAATCTATTG	AAACACTCAA	CTTATCATAA	AGAACTCTTA	AATCTTTAAC	204360
ATAAAAGATG	AAATTTTGCT	CTTTTGAAC	GTGTGACTTT	GAAACTCTAA	AAGCCTTAAA	204420
TCTCATTTTA	CTTGAAGCAA	GAGGATAATT	TGGAACATCG	TCTTTAATAA	TTCTGGATGA	204480

TATATTAGGA	ATATAGTTAG	ATTGACCA	AATTAAATCA	GCCCACCCTT	AAACTTTAA	204540
AGTACCCATA	GAATAAGCAT	ATTCCATGCC	ATTCATATCT	TCAAATAAAA	CCTCAAGATC	204600
TATCTCATAC	CCTAAACTAT	AAACAGATAC	CTTAATTTCT	TTCATGGTTT	TAATGTTATC	204660
AATAAGACCT	TTGCCTAAAA	ATTGATTGCC	ACTTTCCCCT	GAATAAAAAG	GAATTTTAAA	204720
TGGTGGCATA	ATCATAGCAG	ATGATTGAGA	ATAGCTTTGA	AACAAAACCTC	TTACCCCTAA	204780
AATAGTATCA	CCTGCGTACC	TTTTTGACTC	ACTCTTAACA	ACAGCGGGCG	CAACAACCTGA	204840
ATTTTTAACG	TAAGCCTGCA	ACCTTGCAGA	AGGAGTAAGT	AAAACGCTCC	AATTATTTAT	204900
CCCAAGATCT	ACAACCATAT	CTTCCGGCTT	AACAATACCA	GAAGCGCCCG	AATATACATA	204960
ATCAACATAA	TTTGTAAGAT	CAAGTCTAGT	TGAACTTGGA	TCTCTTGCAA	GCTCGGCAAA	205020
ATCTAAAACT	AATTCTCCAG	GCTCTGCCCT	TTTAGAACCC	TCTGCTAATC	CATCAGTCTC	205080
TTGAGCAAAA	AGAACAGTGG	ATAATAAAAA	AAATAAAATA	CTTTTAGCTT	TCCTTTTCAT	205140
GTAAACCAAC	TCCTTATATA	TATAAATCTT	TTATAAAATT	TTCGTTTTTT	AATTATTTTT	205200
ATTAGTAAAA	TTTAATTTAT	CAACTTTGCT	CCTTTAAGAT	ACACCCTTCA	ACAAGAATAA	205260
CAATCTTTTT	TATCTTACCA	GGATAGTTAA	AGCGCATTTG	TTTTATTAAA	GACATAGCAA	205320
AAAGATTGAC	CTTAACATCA	AAAGATTCAT	TAGATTCATT	ATAATCACCA	TTATTAAAAG	205380
AATCATAAAA	TTCTCTTGAA	AGATTTACAT	AATAAACTCC	ATTCTTTAAA	AAAGAATATA	205440
AAAATCTTGA	ATCACTTAAT	AAAAACCCAA	AAGAAAACCC	TTCATTGCTT	CCTAAAAGAA	205500
AATCTTTTAC	TAAAAGATCT	AAATTATCTT	TCAAATTTTG	TTCATCTCTT	AAATATCTTA	205560
AATTAGCAAC	AAATCCCTTG	CTAGAATGAA	AATAAAAAAC	CTTTTTTGAA	AAAAGATTAT	205620
CATAATTTAA	AAAAACCATA	CAAATAGAAA	ACAAAAAGCT	TACTGCCAAA	GACCCCAAAA	205680
GCACCCTAAT	CATATGCTCT	TTTTTTGAAT	TTAAAAATTT	ATAAATCACT	ACATTATATT	205740
TAAAAAATAT	ATCCGAAATA	TTATTTTTC	TAAAAATTTA	TAAATTCCAT	TAAAGCTTTA	205800
AGTATTAGAA	TATTAAACTT	GCTCATGTAA	TTATAATCCA	AAATTAATTT	AGCATCTAAA	205860
ATATTGGATA	AAAACCCCAT	TTCAATCAAC	ACAGCAGGCA	TACTGCTGTT	TTTTATTACA	205920
AACCATTGCT	CTTTTCTGAT	TGGCCTAATA	TTAGTTTCGC	TTAACTCATT	TTTAAACACT	205980
TTATACAAAA	TTTCAGCCAA	TCTTTTGTAT	TCATATTTAT	ATTTAATATC	TAGTATATCA	206040
TTAAGCTCGC	TTAAGTATCT	ATTACCTTTA	ATATCATATC	CCTTAAAATC	TTTAATAACT	206100
TCTCTTTTTG	AATCCTTAGG	AAGATACCAA	AACTCAACTC	CTCTAGCTTC	ACCGTTTGGA	206160
GCATCATTAG	CATGTATAGA	TAAAAATATA	ACATTATTGG	GGAAATTTGG	CTTTATTGCA	206220
TTTGCAAATT	CCGACCGTTC	TTTTAAAGTT	AAATAACAT	CATTTATACG	AGTTAACAAA	206280

ATATTTTTAT	TTACAAAATA	ATTACTTAAA	ATTTTAGACA	AATATATAGA	ATAGGTTAAT	206340
GCAAAATCTT	TTTCCTGAAG	CACAACGTCA	TAACCATTTA	TCTTTAAAGT	CACAACAGCA	206400
CCAGTATCAT	GCCCCCCATG	TCCAGGATCA	ATGATTATTG	AAGTAATTCT	GGGTTTATTA	206460
TAGTCTTTAA	GAGAACTAAA	ATAATTTTCA	ATTTGTTTTA	ATACCTTTTG	ACTTATTAAA	206520
ATTTCTCCAC	GAATGTCAAT	AATTGGGTCT	ACAAACATAT	AATAACCAGA	AGATGTAAGC	206580
GCATATTCAA	AGCCTACCCT	AAACTTCAAA	TATCCCTTAT	CATTTTCAAT	TGTAAAAACA	206640
TCATTTTCAA	TGTTAAATC	AAACCTAAAA	ACATTAGTAT	CAAAAAAATC	AAGAACATTT	206700
AAATAATCGG	GGGTCTTAGA	ATACAAGCTT	AAATATGAAA	ACAAAATCAA	ATCAATCAAT	206760
AATATCATTT	TCCCAAAGCT	CAATGGCACT	CTTTAAGTCT	CCTTTGAATC	TTAAGCTATT	206820
TTTAGTAATC	TCACCTTTTA	TCTCCTTAAA	CTCTTTTTTG	AAGAGAAAAG	GATTAATTTT	206880
GTATTTTAAA	CAAATTTTGC	AAAACCTTAA	AAAAGAAAAA	ACAACGCCAT	TTCCAACAAA	206940
AATTGGGACA	TAATTCTTAA	GCAAAAAAAT	CAAAAAACTT	TTGCTTTTGG	TTAAATTTTC	207000
TGGCGAATTT	AATGCAGTAT	ACTGGATTCT	CAGCTTTCTA	TAAATATAAA	CATGCTCCCC	207060
AAAATAAATA	GCTCTTAGTC	CAAAATCTAT	TCTTTGAAAA	TATTCATTTT	GAATTCCTGC	207120
GTCAAAACCT	CCAAGCTGTA	AAAATTTCTC	TTTAGAATAA	AGTCCGCAAT	AATCCATGGT	207180
AATCAAAGTT	TTTTCATAGT	CTTCTCAGA	ATTTACTAAA	ATTACCTTAA	ACTTTTGCTT	207240
TTTATCTATG	CTGGGAAGAA	AAATTGAAGG	AATCATTTCC	TCTTCTTTAT	CAAAAAACTC	207300
CCCACCAACA	AGAAGAACAT	TTTTTTTTTAC	TATTTTCATCG	AATATATTTG	GAATCCAAAA	207360
GGGATTTAAC	AAGTACATAT	CACTTTGCAA	AACAAAAACA	AAATCACAGC	TGGATTCTTT	207420
CATTGCTAAA	TTAACCTTTT	CCCCAGAATT	CAAATCGTCA	GAAAGTAAAA	TAAATTTTAA	207480
CTTACCATAA	CTTCTGAAA	TAAACTGCAA	AGAACTTCTA	TTGCTCTGTT	TTTCAATTGA	207540
AATTATTTCT	CTTATAAAGT	CAAAATTTGA	TAAAAATTCA	AACAAATCTT	CTCTAAAAAT	207600
TTTTGTTCCT	CTGCTTAATA	TTACAAAAGA	AATTCCAAAA	GAAGATTTTT	GTGAATAATT	207660
ATTTTITAGAT	TGAATAACAG	TATATGAATA	GCCACTACCT	GGAAGACGCA	TAAATTACTT	207720
TTAAAATCCT	TATAATTAAA	TTATAATAAT	CATATGTTAC	ATAATACAAT	GCTAATTGCA	207780
AGAATAATGA	ATATTAATAC	ATTATTCTAC	GGCATGATCA	TTATCATTTT	TGCACTCATT	207840
TCTTGCAATC	ATAAGAATAT	ACAGTACGAC	AAGAGAATTA	AAAAATTTTT	AGATAAAAAAC	207900
AAAATTGAAT	ATAAAATAGA	CTCAGAAAAT	GACTTTATAG	CATTTAAAGA	TATAACAAT	207960
AACGAAAAAG	AAGAAGTAAT	CATCAGATCA	AGACTAAACT	CATATAAAAA	TTCAAAGATA	208020

AGAGAAATAT	TTGGAATTGT	TAAAGTATTT	GATATAAACA	CACCAAAAAT	TAAAGAAATA	208080
TCTGACTCGC	TTATGAGCGA	TAGTTATAAT	AACAGAGTAT	TTGGATCGTG	GGAGATTATT	208140
CATAATGCAG	AAAGAGGAAT	CAACTCTTTG	GTATATATTG	TAAAAGCAGA	AGAATTTGCA	208200
AATGATACAT	TTTTGCTTGA	TGCAATTGAT	GAGATTGCCT	CAACAATAAG	TATTTTCAAA	208260
AAAATAATAA	CAACCAACAA	CGAAAACATT	GATAATAATG	AAGAAAATAA	CAATACAAAT	208320
GAATCAAATG	AACAGCCCAC	CTTAAAGCAA	GAAAAAACAA	ATTCAACAAA	AGAATCTAAT	208380
AACGAACTTA	AAGAAGATCA	AATAGAAGAA	GAACCTCAAG	AAATCAAAGC	CCAATAATTT	208440
CAAAATCATT	CTACTAATAA	AGAATTAACA	TCAAAGCAAA	AATGAACCTT	GTCACCTATT	208500
TTTATTTGAA	ATTTAGAAAA	TGAGCCCCTT	GGAAGCTCAA	GAGCATACCT	TACCTTATAA	208560
AGAGAATTAA	CATTTGCCCT	AGAGTACGGC	TCTAAATCAT	AAATCTCTTT	AATAATTCCA	208620
CTTGAATCAA	TATAAGCTAT	TTCAAGCAGC	AAAGGTGTAT	TTCCATCCA	AAAAGACAAA	208680
TTTTGATCTT	TTTTAAAAAC	AAAAAGCATT	CCATTGCCAT	ATTCAACTTT	TTGAGCACCC	208740
ATGTAACCTT	TTGCCCTATC	AAGCTCATT	GATGCTATTT	TTACAAAAAA	CTTAACCCCA	208800
TTTATCATAA	TTTCTTTATC	GTAAAAATGA	TCAGCAAAAG	ATAAAAAAGA	CATCGACAAA	208860
ACTAAAAACA	AAAACCGTTT	TAAAATTTTT	TTCAATTATC	AGCCTTATTA	AAAATCATTT	208920
ATTATAATTT	GAAATATAAG	ATTTTAAAGT	AATTCTTAAA	ATATTTTTTAT	TTAAAACAAT	208980
AATAGAATCA	CCAAGATCCC	ATAAATAATT	CAAAGGCGTT	ACTAGTGTAG	CCTCGCCAAA	209040
CTTTTTCTTT	ACCTCAATAA	AAATTGCATA	CTGAGATATT	AATTTTTTTAT	CAAAAACAAA	209100
AGTGTACGAA	AACATTTTAT	TTTTGTCAAA	ATTAAAAACA	GCATATTTTA	AATGGGTTTT	209160
AGCAATAAAT	TGTAAAGTCT	GCTTCCTAGC	ATAAGGAAAT	TCAACTTCTT	CAACATCATA	209220
ATAAAAAATA	GTGCTCTTGA	GAACATTCTC	TTTAACAGAA	TCCATGCTAG	AACCCAGAAC	209280
AAAAGATGCA	AAAATAGAAA	ATAAATAAAA	AACTGTAATA	CCCATACAAA	GCCTTTAGCC	209340
ATAACAATA	TAAAGAAAAA	GTATAACAAA	TAAAAAACTT	GAAAATAAAT	AAAGCAAAAT	209400
ATCAATATTA	ATCCCAAAAT	GGAAAAGCCT	TACAAAGGTC	TTAGAAGCTG	CTCTATCACT	209460
TACCATTCTT	GCTCTTTCTT	TTCTTTTAAA	GGGCTTTTCT	TTAAGTTCTT	GAAAAGGACT	209520
GTAGTGACAA	TTAGGACAAC	CTTCTCCAAA	AGCAGAAACT	GGACCTACAT	GCCTACAATT	209580
TGGACATTCTG	ATATCACCAA	GCTTAGCAGC	ACAGTTGGGA	CAAACAGATC	GATTAAGTCC	209640
AACTTTTTCTG	GCACATTGCT	CGCAAAAAAC	TTCAAAATTT	ACCTTTGCCA	AACGAAAATT	209700
CCTCAAATTT	TACTTAAAAA	TAAGTATTTA	AAATACTATA	TAATTAATTA	TAATAAAAAA	209760
TATATGAATA	TTACATATTT	AAGGAATACT	AAAACATGAA	ATCGGGATTT	GCAGCAATAC	209820

TTGGTAGACC	ATCAACTGGA	AAATCTACCC	TTTTAAATTC	AATATGCGGA	CATAAAATAT	209880
CAATAATATC	CCCTATTCCG	CAAACAATA	GAAATAATAT	AAAAGGAATC	TTTACGGACG	209940
ACAGAGGACA	AATTATTTTT	ATAGACACAC	CGGGATTTC	TCTGAGTAA	AAAAAGTTTA	210000
ATATTGCAAT	GATGAAAAAT	ATCCACTCTT	CAATAGGAGA	AGTTGAACTC	ATTTTATACA	210060
TAATAGACAT	TCAAGACAAA	CCTGGAGAAG	AAGAAAATAA	AATGTTAGAA	ATAATTAAAA	210120
ACTCTAAAAT	TAAATTTTTA	GTAATACTTA	ATAAAATTGA	CCTTAAAAAC	ACAAAAATAA	210180
AAGAAATAAC	GCAATTTCTA	AAAGAAAAAG	GAATAGAAGA	TAGTAATATA	ATTAAATAT	210240
CTGCTGAAAA	AAAAATTAAC	ACAGAAGAAC	TAAAAAATAA	AATTTATGAA	AATTTTTCAG	210300
AAGGCCCACT	TTATTATCCA	CAAGAATACT	ACACCGATCA	AGAAATAAAT	TTTAGAATTA	210360
GTGAAATAAT	AAGGGAAAAA	GCTATTGAAA	ACCTAAAAGA	AGAACTCCCC	TATTCTTTGT	210420
ATGTGGATAT	TGATACCTTA	GAAAATAAAA	AAGGAAGTCT	TTTATCAGA	GCAAATATTT	210480
TTGTAGCCAA	TGAAAGTCAA	AAAGGAATAA	TTGTAGGAAA	AAACGGAAAA	GAAATAAAAT	210540
CAATAGGAGA	AAGGGCAAGA	AAAACAATTG	CAAAAATTTT	TGAAACAAAA	TGCAACCTAT	210600
TC'TTACAGGT	AAAAC'TTAA	AAAAATTGGA	ACAAAGAAGA	TAAGCTAATA	AAAAGACTTA	210660
TAAATTAACA	AACATTAAAC	TGCATTTTTT	TAAATCTTGT	AAACTTGAAA	AACAAAATGC	210720
TAAAATTTAC	CTAAATTTAA	ATTAGGAATA	AAATGTGAAA	ACAGCACACT	GGGCAGATTT	210780
TTACGCAGAA	AAAATAAAAA	AAGAAAAAGG	TCCAAAAAAC	TTATACACAG	TAGCATCGGG	210840
AATTACTCCA	TCTGGAAGTG	TGCACATTGG	CAATTTTAGA	GAAGTTATTT	CGGTAGACCT	210900
TGTAGCAAGA	GACTAAGAG	ACTCTGGATC	AAAAGTAAGG	TTTATTTATT	CTTGGGATAA	210960
TTACGACGTA	TTTCGAAAAG	TTCCCAAAAA	TATGCCAGAA	CAAGAACTTC	TTACAACTTA	211020
TTTAAGACAA	GCAATAACAA	GGGTCCCTGA	CACAAGAAGC	CACAAAACAA	GTTATGCAAG	211080
GGCTAATGAA	ATTGAATTTG	AAAAATATCT	GCCTGTAGTT	GGGATCAATC	CTGAATTCAT	211140
CGACCAAAGC	AAACAATATA	CCAGCAACGC	TTATGCAAGC	CAAATAAAAT	TTGCACTTGA	211200
TCATAAAAAA	GAAGTGTCTG	AAGCATTAATA	CGAATACAGA	ACCTCAAAGC	TTGAAGAAAA	211260
TTGGTATCCA	ATCAGTGTAT	TTTGTACAAA	ATGCAATAGA	GACACAACAA	CTGTAAATAA	211320
TTATGACAAT	CATTACTCTG	TTGAGTATTC	ATGTGAATGT	GGAAATCAAG	AATCTCTAGA	211380
CATAAGAACC	ACATGGGCCA	TTAAACTTCC	TTGGAGAATA	GATTGGCCTA	TGAGATGGAA	211440
ATATGAAAAA	GTTGACTTTG	AGCCTGCAGG	AAAAGACCAC	CACAGCAGTG	GCGGCAGTTT	211500
TGATACATCT	AAAAATATTG	TAAAAATTTT	TCAAGGTAGC	CCTCCTGTAA	CATTTCAATA	211560